

Université de Montréal

**Découverte d'inhibiteurs de la dihydrofolate réductase R67  
impliquée dans la résistance au triméthoprim.**

par

Dominic Bastien

Département de biochimie

Faculté de médecine

Mémoire présenté à la Faculté des études supérieures

en vue de l'obtention du grade de *M.Sc.*

en biochimie

Août 2012

© Dominic Bastien, 2012

Université de Montréal  
Faculté des études supérieures

Ce mémoire intitulé :

Découverte d'inhibiteurs de la dihydrofolate réductase R67 impliquée dans la résistance au triméthoprim.

Présenté par :

Dominic Bastien

a été évalué par un jury composé des personnes suivantes :

Dr. James Omichinski

Dr. Anne Marinier

Dr. Joelle Pelletier

## Résumé

Le triméthoprim (TMP) est un antibiotique communément utilisé depuis les années 60. Le TMP est un inhibiteur de la dihydrofolate réductase (DHFR) bactérienne chromosomale. Cette enzyme est responsable de la réduction du dihydrofolate (DHF) en tétrahydrofolate (THF) chez les bactéries, qui lui, est essentiel à la synthèse des purines et ainsi, à la prolifération cellulaire. La résistance bactérienne au TMP est documentée depuis plus de 30 ans. Une des causes de cette résistance provient du fait que certaines souches bactériennes expriment une DHFR plasmidique, la DHFR R67. La DHFR R67 n'est pas affectée par le TMP, et peut ainsi remplacer la DHFR chromosomale lorsque celle-ci est inhibée par le TMP. À ce jour, aucun inhibiteur spécifique de la DHFR R67 est connu. En découvrant des inhibiteurs contre la DHFR R67, il serait possible de lever la résistance au TMP que la DHFR R67 confère aux bactéries.

Afin de découvrir des inhibiteurs de DHFR R67, les approches de design à base de fragments et de criblage virtuel ont été choisies. L'approche de design à base de fragments a permis d'identifier sept composés simples et de faible poids moléculaire (fragments) inhibant faiblement la DHFR R67. À partir de ces fragments, des composés plus complexes et symétriques, inhibant la DHFR R67 dans l'ordre du micromolaire, ont été élaborés. Des études cinétiques ont montré que ces inhibiteurs sont compétitifs et qu'au moins deux molécules se lient simultanément dans le site actif de la DHFR R67. L'étude d'analogues des inhibiteurs micromolaires de la DHFR R67 a permis de déterminer que la présence de groupements carboxylate, benzimidazole et que la longueur des molécules influencent la puissance des inhibiteurs. Une étude par arrimage moléculaire, appuyée par les résultats *in vitro*, a permis d'élaborer un modèle qui suggère que les résidus Lys32, Gln67 et Ile68 seraient impliqués dans la liaison avec les inhibiteurs. Le criblage virtuel de la librairie de 80 000 composés de Maybridge avec le logiciel Moldock, et les essais d'inhibition *in vitro* des meilleurs candidats, a permis d'identifier quatre inhibiteurs micromolaires appartenant à des familles distinctes des composés précédemment identifiés. Un second criblage virtuel, d'une banque de 6 millions de composés, a permis d'identifier trois inhibiteurs micromolaires toujours distincts. Ces résultats offrent la base à partir de laquelle il sera possible de développer

des composés plus efficaces et possédant des propriétés pharmacologiquement acceptables dans le but de développer un antibiotique pouvant lever la résistance au TMP conféré par la DHFR R67.

Mots clés : design à base de fragments, criblage virtuel, arrimage moléculaire, découverte de médicaments, dihydrofolate réductase R67, résistance bactérienne, triméthoprim, inhibition enzymatique.

## Abstract

Trimethoprim (TMP) is a common antibiotic which is used since the 60's. TMP is an inhibitor of the bacterial chromosomal dihydrofolate reductase (DHFR). This enzyme catalyses the reduction of the dihydrofolate (DHF) to tetrahydrofolate (THF) which is essential to the biosynthesis of purines thus to cellular proliferation. Bacterial TMP resistance is documented since about 30 years. One of the cause of this resistance comes from the fact that certain bacteria express a plasmidic DHFR, the R67 DHFR, which confers TMP resistance. The R67 DHFR is not inhibited by TMP and can replace the chromosomal DHFR when the latter is inhibited by TMP. The discovery of R67 DHFR inhibitors would allow to break the trimethoprim resistance granted by R67 DHFR.

In order to discover R67 DHFR inhibitors, fragment based design and virtual screening approaches were selected. By fragment based design, seven simple compounds with a low molecular mass which inhibited weakly R67 DHFR (fragments) were identified. From these fragments, more complex and symmetrical compounds inhibiting R67 DHFR in the micromolar range were identified. Kinetic studies showed these inhibitors were competitive and at least two molecules bind simultaneously to the active site of the R67 DHFR. Test of the micromolar inhibitors analog showed that the presence of carboxylate, benzimidazole and the length of the molecule all have an effect on the potency of the inhibitors. Molecular docking of the inhibitors, supported by *in vitro* data, were used to develop a model which suggest that residue like Lys32, Gln67 and Ile68 would be involved in the binding of the inhibitors to the R67 DHFR. Virtual screening of the 80 000 compound Maybridge library with Moldock software, followed by *in vitro* test of the best candidate, identified four micromolar inhibitors which are chemically distinct from the inhibitor beforehand identified. A second virtual screening of a 6 million compounds bank identified three micromolar inhibitors which are also distinct from the inhibitor beforehand identified.

These results offer a basis which will allow further development of more potent inhibitors with more acceptable pharmacologic properties in order to develop an antibiotic which would break the TMP resistance granted by the R67 DHFR.

Key words : fragment based design, virtual screening, molecular docking, drug discovery, R67 dihydrofolate reductase, bacterial resistance, trimethoprim, enzymatic inhibition

## Table des matières

<b>Chapitre 1</b> - Introduction.....	15
<b>Section 1.0</b> - Avant-propos.....	15
<b>Section 1.1</b> - Le triméthoprim.....	16
Section 1.1.1 -Le triméthoprim : un antifolate.....	17
Section 1.1.2 - Les dihydrofolates réductases.....	19
Section 1.1.3 - Résistance aux inhibiteurs de dihydrofolate réductase.....	21
<b>Section 1.2</b> - La dihydrofolate réductase R67.....	24
Section 1.2.1 - Site actif de la dihydrofolate réductase R67.....	28
Section 1.2.2 - Inhibiteurs connus de la dihydrofolate réductase R67.....	31
<b>Section 1.3</b> - Du gène au médicament.....	33
Section 1.3.1 -Découverte de la cible.....	34
Section 1.3.2 - Identification de composés actifs.....	39
Section 1.3.3 - Détermination et caractérisation de la tête de série.....	46
Section 1.3.4 - Optimisation de la tête de série et tests précliniques.....	49
Section 1.3.5 - Essais cliniques.....	49
<b>Section 1.4</b> - Identification de composés actifs : inhibiteurs.....	51
Section 1.4.1 -Design à base de fragments.....	51
Section 1.4.2 - Arrimage moléculaire.....	55
<b>Section 1.5</b> - Profilage des inhibiteurs de la DHFR R67.....	60
Section 1.5.1 -Évaluation de la puissance par dose-réponse.....	61
Section 1.5.2 - Sélectivité par rapport à la dihydrofolate réductase humaine.....	62
Section 1.5.3 - Efficacité de liaison.....	63
Section 1.5.4 - Mode d'inhibition.....	64

<b>Chapitre 2</b> - Design de l'essai automatisé.....	77
<b>Section 2.1</b> - Mesure de l'activité de la dihydrofolate réductase R67.....	77
<b>Section 2.2</b> - Détermination des conditions de criblage.....	79
Section 2.2.1 - Solubilité des composés.....	79
Section 2.2.2 - Absorbance des composés.....	81
Section 2.2.3 - pH du milieu réactionnel.....	82
<b>Section 2.3</b> - Fiabilité de l'essai.....	85
 <b>Chapitre 3</b> - Design à base de fragments.....	86
<b>Article 1.</b> Fragment-based design of symmetrical bis-benzimidazoles as selective inhibitors of the trimethoprim-resistant, Type II R67 dihydrofolate reductase.....	86
<b>Section 3.1</b> - Présentation de l'article et mise en contexte.....	87
ABSTRACT.....	90
INTRODUCTION.....	91
MATERIALS AND METHODS.....	94
RESULTS AND DISCUSSION.....	99
CONCLUSION.....	114
SUPPLEMENTARY MATERIALS.....	122
 <b>Chapitre 4</b> - Relation structure activité.....	137
<b>Section 4.1</b> - Orientation des substituants du phényle.....	138
<b>Section 4.2</b> - Effet de la longueur de l'espaceur.....	139
<b>Section 4.3</b> - Effet du groupement benzimidazole.....	139



<b>Section 4.4 - Design de nouveaux inhibiteurs.....</b>	<b>140</b>
Section 4.4.1 - Modifications des extrémités.....	141
Section 4.4.2 - Modifications des cycles.....	142
Section 4.4.3 - Connexions de deux molécules d'inhibiteurs.....	142
Section 4.4.4 - Favoriser une conformation linéaire ou repliée.....	143
Section 4.4.5 - Modifications des espaceurs.....	144
<b>Section 4.5 - Arrimage moléculaire des composés 23 à 34.....</b>	<b>145</b>
 <b>Chapitre 5 - Arrimage moléculaire</b>	
<b>Section 5.1 - Criblage de la librairie virtuelle de Maybridge.....</b>	<b>148</b>
Section 5.1.1 - Choix du logiciel Moldock.....	149
Section 5.1.2 - Résultats du criblage virtuel.....	150
Section 5.1.3 - Test <i>in vitro</i> de dix des composés Maybridge identifiés <i>in silico</i> .....	152
<b>Section 5.2 - Criblage de la librairie du CNRC.....</b>	<b>154</b>
Section 5.2.1 - Résultats du criblage virtuel.....	155
Section 5.2.2 - Test <i>in vitro</i> des composés identifiés par le criblage virtuel du CNRC.....	158
 <b>Chapitre 6 - Conclusion.....</b>	
<b>Annexe I - Matériel et Méthodes.....</b>	<b>168</b>

## Liste des tableaux

### Chapitre 1

<b>Table 1.1</b> Caractéristique de la DHFR microbienne chromosomale, de la DHFR R67 et de la DHFR humaine.....	23
---	----

### Chapitre 2

<b>Table 2.1</b> Absorbance pour des volumes de 100 $\mu$ L et 200 $\mu$ L du composé 15.....	82
---	----

<b>Table 2.2</b> pH de la solution après l'ajout de 10 mM des composés types criblés.....	84
---	----

### Chapitre 3

<b>Table 3.1</b> Inhibition of R67 DHFR and hDHFR by fragments 1-7.....	100
---	-----

<b>Table 3.2</b> Inhibition of R67 DHFR and hDHFR by symmetrical compounds 8 and 9.....	103
---	-----

<b>Table 3.3</b> Docking of NADPH and 8, 8a, 8b and 9 to R67 DHFR.....	107
--	-----

<b>Table 3.4</b> Cytotoxicity of 8, 8a, 8b and 9 toward 3T6 fibroblast cells.....	113
---	-----

<b>Table S1</b> List of the compounds screened against R67 DHFR.....	123
--	-----

### Chapitre 4

<b>Table 4.1</b> Comparaison du IC50 des composés 9 et 10.....	138
--	-----

<b>Table 4.</b> Arrimage moléculaire des composés 23 à 34.....	146
--	-----

## Chapitre 5

<b>Table 5.1</b> Soixante composés ayant obtenu le meilleur score suite au criblage de la librairie Maybridge.....	150
<b>Table 5.2</b> Composé choisit pour les tests <i>in vitro</i> .....	151
<b>Table 5.3</b> Composés Maybridge offrant une inhibition de la DHFR R67.....	153
<b>Table 5.4</b> Composés provenant du criblage virtuel du CNRC choisis pour les tests <i>in vitro</i> .....	158

# Liste des figures

## Chapitre 1

<b>Figure 1.1</b> Inhibiteurs de DHFR utilisés présentement en clinique.....	17
<b>Figure 1.2</b> Structure du DHF, NADPH et de la réaction catalysée par la dihydrofolate réductase.....	19
<b>Figure 1.3</b> Implication du THF dans la biosynthèse des purines et des désoxy thymidines monophosphates.....	20
<b>Figure 1.4</b> Concept schématisé d'inhibiteurs de DHFR R67 pour contrer la résistance au TMP conféré par la DHFR R67.....	24
<b>Figure 1.5</b> Structure quaternaire de la DHFR R67.....	25
<b>Figure 1.6</b> Mécanisme de catalyse proposé pour la DHFR R67.....	26
<b>Figure 1.7</b> Influence du pH sur l'activité de la DHFR R67.....	27
<b>Figure 1.8</b> Implication des résidus clés dans la liaison du DHF et du NADPH dans le site actif de la DHFR R67.....	28
<b>Figure 1.9</b> Inhibiteurs connus de la DHFR R67.....	31
<b>Figure 1.10</b> Vue d'ensemble du processus du développement d'un médicament.....	34
<b>Figure 1.11</b> Vue d'ensemble du processus de découverte d'une nouvelle cible.....	35
<b>Figure 1.12</b> Évaluation et optimisation parallèle de l'activité et des propriétés médicamenteuses des composés actifs.....	46
<b>Figure 1.13</b> Courbe dose-réponse.....	47
<b>Figure 1.14</b> Processus typique de design à base de fragments.....	52
<b>Figure 1.15</b> Processus typique de criblage par arrimage moléculaire.....	55
<b>Figure 1.16</b> Schématisation du fonctionnement d'un algorithme génétique.....	57
<b>Figure 1.17</b> Graphique d'une courbe de $IC_{50}$ typique.....	61
<b>Figure 1.18</b> Graphique Dixon typique pour les 3 modes d'inhibition réversibles.....	64

## Chapitre 2

<b>Figure 2.1</b> Mesure de l'activité de la DHFR R67 par spectrophotométrie.....	78
<b>Figure 2.2</b> Station robotisée utilisée lors du criblage d'inhibiteur contre la DHFR R67.....	78
<b>Figure 3.3</b> Activité des DHFR en présence de différents co-solvants organiques.....	80
<b>Figure 3.4</b> Limite de détection de la lectrice à plaque DTX880.....	81
<b>Figure 3.5</b> Titrage acido-basique de l'activité de la DHFR R67 par des bases faibles...	83

## Chapitre 3

<b>Figure 3.1</b> Structures of the weak inhibitor fragments and their inactive analogs.....	99
<b>Figure 3.2</b> Structures of the R67 DHFR inhibitors and their inactive analogs.....	102
<b>Figure 3.3</b> Dixon plot for determination of the type of inhibition of R67 DHFR by 8, relative to the NADPH cofactor.....	105
<b>Figure 3.4</b> Docking of 8, 8a, 8b and 9 R67 DHFR using Molegro Virtual Docker.....	109
<b>Figure 3.5</b> Hill plot for determination of the stoichiometry of binding of compound 9 to R67 DHFR.....	112
<b>Figure S1</b> IC <sub>50</sub> plot of 8 and 9.....	134
<b>Figure S2</b> Dixon plot for determination of the type of inhibition of R67 DHFR by 8, relative to the DHF substrate.....	135
<b>Figure S3</b> Docking of compound 9 on R67 DHFR using AutodockVina.....	136

## Chapitre 4

<b>Figure 4.1</b> Analogue des inhibiteurs 8 et 9 qui ont été testés sur la DHFR R67.....	137
<b>Figure 4.2</b> Valeur des IC <sub>50</sub> des composés 11 à 16.....	139
<b>Figure 4.3</b> Valeur des IC <sub>50</sub> des composés 17 à 22.....	140
<b>Figure 4.4</b> Modifications proposées sur les extrémité des inhibiteurs.....	141

<b>Figure 4.5</b> Modification pour combler l'absence du fragment benzimidazole.....	142
<b>Figure 4.6</b> Connexions de bras de molécules inhibitrices de la DHFR R67.....	143
<b>Figure 4.7</b> Composés qui favorisent la conformation linéaire (30) ou repliée (31).....	144
<b>Figure 4.8</b> Modification de l'espaceur.....	145

## Chapitre 5

<b>Figure 5.1</b> Forme holoenzyme (A) et apoenzyme (B) de la DHFR R67.....	155
<b>Figure 5.2</b> Distribution selon les scores (A et B) et la masse molaire (C) suite au criblage virtuel de la librairie du CNRC.....	157
<b>Figure 5.3</b> Tests <i>in vitro</i> des 21 composés provenant du criblage virtuel du CNRC.....	161
<b>Figure 5.4</b> Structure des 3 inhibiteurs découverts provenant du criblage du CNRC....	162

# **Chapitre 1 - Introduction**

## **Section 1.0 - Avant-propos**

Le développement de médicaments est une des entreprises les plus ambitieuses et coûteuses qu'on retrouve en recherche. Évidemment, ce processus dépasse largement les objectifs du présent projet de maîtrise. Malgré plusieurs étapes qui n'ont pas été abordées lors de ce présent travail, toutes les étapes du développement du médicament seront discutées afin d'être en mesure de bien situer et comprendre les travaux qui ont été accomplis. L'accent sera mis sur les étapes que couvrent cette recherche, c'est-à-dire, la découverte de molécules actives contre la dihydrofolate réductase R67 (DHFR R67) selon le design par fragment et par arrimage moléculaire. Par la suite, il sera question de la caractérisation de ces inhibiteurs par la détermination de leur puissance et leur sélectivité envers la DHFR R67 par rapport à son équivalent chez l'humain, la dihydrofolate réductase humaine (hDHFR). Leurs mécanismes d'inhibition et leurs interactions avec la DHFR R67 seront ensuite étudiés. Bien que l'objectif principal de ce mémoire soit de rapporter les avancées faites au niveau du développement d'inhibiteurs contre la DHFR R67, l'objectif secondaire de ce mémoire se trouve à initier des personnes, inexpérimentées ou hors de ce domaine, au développement du médicament.

## Section 1.1 - Le triméthoprim

La résistance aux antibiotiques est un problème de santé publique majeur. Le taux de résistance aux antibiotiques a atteint des niveaux alarmants<sup>1</sup>. La résistance aux antibiotiques a un coût humain et économique important. En 2007, seulement en Europe, environ 400 000 infections par des bactéries multi-résistantes ont été répertoriées<sup>2</sup>. Plus de 25 000 morts ont été attribuées à ces infections<sup>2</sup>. Aux États-Unis, les coûts engendrés par la résistance aux antibiotiques se situent à près de 35 milliards de dollars par année<sup>2</sup>. La consommation massive d'antibiotiques est reconnue comme étant la principale cause de la hausse de résistance<sup>3</sup>. En plus de leur grande utilisation dans le domaine médical, les antibiotiques sont utilisés dans le domaine vétérinaire; environ 50% des antibiotiques utilisés le sont actuellement dans des élevages de toutes sortes<sup>2</sup>. Ainsi, des efforts internationaux sont requis pour combattre ce phénomène<sup>1</sup>. Parmi les divers moyens proposés pour contrer la propagation de la résistance aux antibiotiques, la conception de nouvelles molécules ou de molécules analogues aux antibiotiques existants, qui pourraient échapper aux mécanismes de résistance, est recommandée<sup>2</sup>.

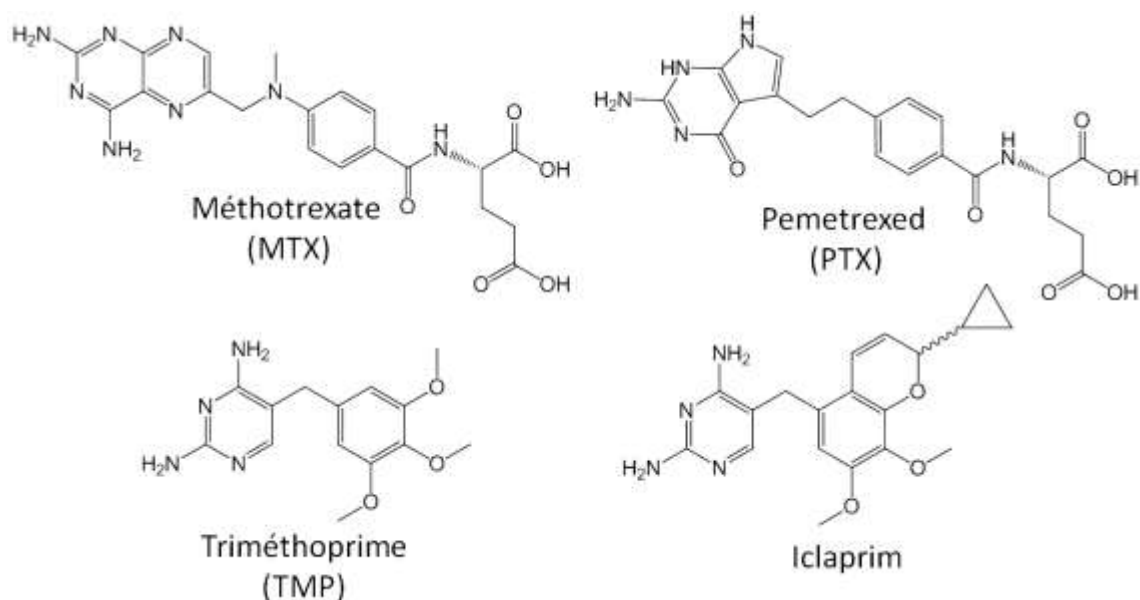
Le triméthoprim (TMP) est un antibiotique commun qui est utilisé depuis le début des années '60. C'est l'antibiotique de première ligne dans le traitement des infections urinaires; seulement en Ontario, environ 4000 prescriptions de TMP sont émises, chaque semaine<sup>4</sup>. Il est aussi utilisé dans le traitement d'infections respiratoires<sup>4</sup>, entériques et plus récemment, dans le traitement de la tuberculose<sup>5</sup>. En plus de ses applications médicales chez les humains, le TMP est utilisé dans l'élevage de saumons<sup>6</sup>, crevettes<sup>7</sup> et bovins<sup>8</sup>. Le rejet massif de TMP dans l'environnement est inquiétant du fait que cette pression sélective provoque l'apparition de bactéries résistantes au TMP. Ces bactéries peuvent par la suite transférer leur résistance à des pathogènes humains par transfert horizontal<sup>9</sup>. Par exemple, il est reconnu que la bactérie *Vibrio cholerae* responsable de l'épidémie de choléra de 1992 en Amérique latine avait acquis sa résistance aux antibiotiques en étant en contact avec des bactéries sélectionnées par l'usage massif d'antibiotiques dans les fermes d'élevages de crevettes en Ecuador<sup>10</sup>. Bien que l'utilisation des antibiotiques dans les élevages est fortement réglementée dans les



pays industrialisés, des pays émergents tels que le Chili et la Chine continuent d'utiliser massivement les antibiotiques dans leurs fermes d'élevage<sup>11</sup>. Ainsi, la grande utilisation du TMP chez les humains et son rejet massif dans l'environnement favorise l'apparition de résistance envers cet antibiotique. Le but de ce projet de recherche est de développer des inhibiteurs contre la DHFR R67 qui est reconnue comme étant une enzyme impliqué dans la résistance au TMP.

### Section 1.1.1- Le triméthoprim : un antifolate

Les antifolates, dont le TMP fait partie, sont une classe de médicaments qui perturbent le chemin métabolique du folate qui est essentiel chez toutes les cellules. À la figure 1.1, on peut voir quelques antifolates qui sont actuellement utilisés en clinique. Ce sont tous des inhibiteurs qui perturbent la voie du folate par l'inhibition de l'enzyme dihydrofolate réductase (DHFR). Ces enzymes sont responsables de la biosynthèse du tétrahydrofolate (THF) qui entre dans la synthèse de précurseurs de l'ADN. De ce fait, le THF est essentiel à la prolifération cellulaire (voir section 1.1.2).



**Figure 1.1 : Quelques inhibiteurs de DHFR utilisés présentement en clinique.** L'iclaprim se trouve en phase clinique III.

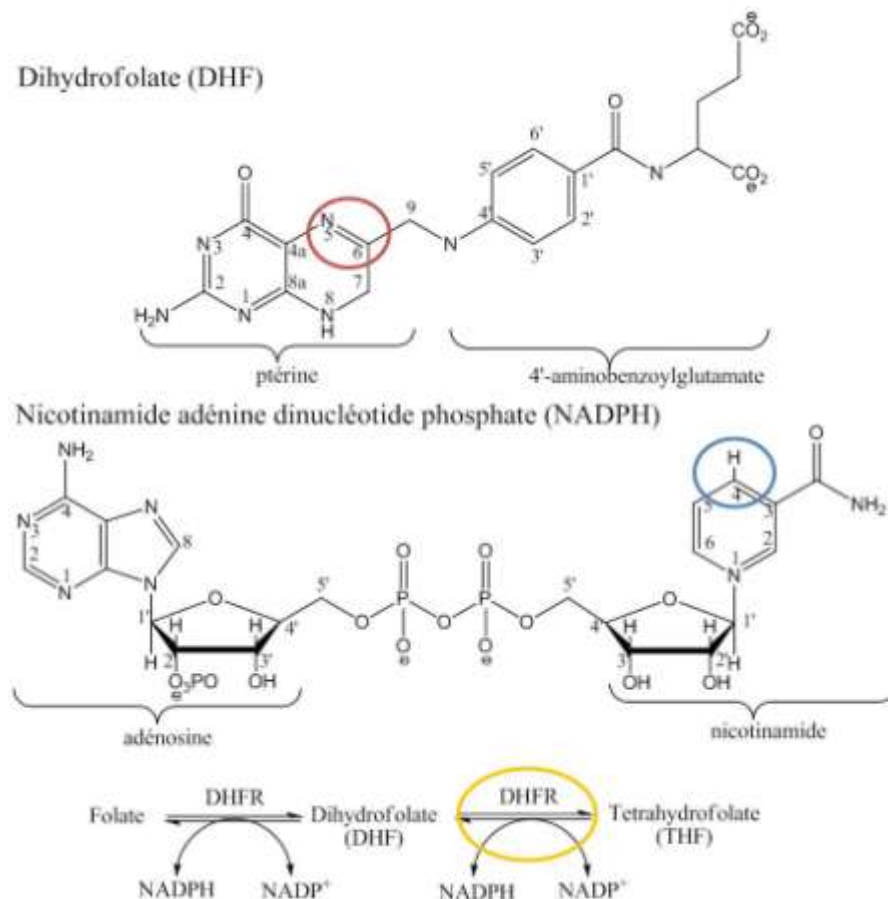
Certains médicaments qui visent les DHFR (figure 1.1) sont utilisés depuis plus de 60 ans<sup>12</sup>. Ils sont utilisés en clinique dans le traitement de maladies à prolifération cellulaire rapide comme le cancer et les infections bactériennes. L'efficacité de ces inhibiteurs provient du fait que les cellules qui prolifèrent rapidement ont un besoin accru en précurseurs de l'ADN afin de répliquer leurs génome. Ainsi, par l'inhibition de la DHFR, on réduit la capacité des cellules à produire du THF ce qui bloque la propagation du cancer et des infections bactériennes.

Le méthotrexate (MTX) a été le premier inhibiteur de DHFR utilisé en clinique. C'est un puissant inhibiteur de la DHFR humaine avec une constante d'inhibition dans les environs de 5 pM<sup>13</sup>. Il était seulement utilisé pour le traitement de la leucémie lymphoblastique aiguë, mais son utilisation s'est élargie. Maintenant, il aide dans le traitement du cancer du sein, des poumons, les cancers lymphatiques ou le cancer de la peau<sup>14</sup>. Le développement d'inhibiteurs contre la DHFR humaine est toujours d'actualité: il y a un peu moins de 10 ans, le pemetrexed (PTX) a été introduit dans le traitement du cancer des poumons<sup>15</sup>.

Pour les infections bactériennes, la DHFR microbienne est aussi une cible intéressante. Le triméthoprim (TMP) est un inhibiteur de la DHFR microbienne avec une constante d'inhibition dans les environs de 20 pM<sup>16</sup>. Il a été introduit sur le marché en 1962 et il est efficace contre les bactéries Gram négatives impliquées dans les infections urinaires, respiratoires, de la peau et entériques<sup>17</sup>. Comme vu dans les paragraphes précédents, le TMP est encore actuellement utilisé en clinique; c'est le médicament de première ligne contre les infections urinaires<sup>18</sup>. Vu la grande efficacité clinique et la faible toxicité du triméthoprim, les industries et les institutions académiques continuent le développement d'inhibiteurs de la DHFR microbienne. Afin de contrer la résistance au TMP, Roche et Arpida ont développé un dérivé du TMP : l'iclaprim<sup>19</sup>. En 2006, l'iclaprim venait de passer avec succès la phase clinique II et entrait en phase clinique III<sup>20</sup>. Récemment, une activité de l'iclaprim contre des bactéries Gram positives a été démontrée<sup>21</sup>.

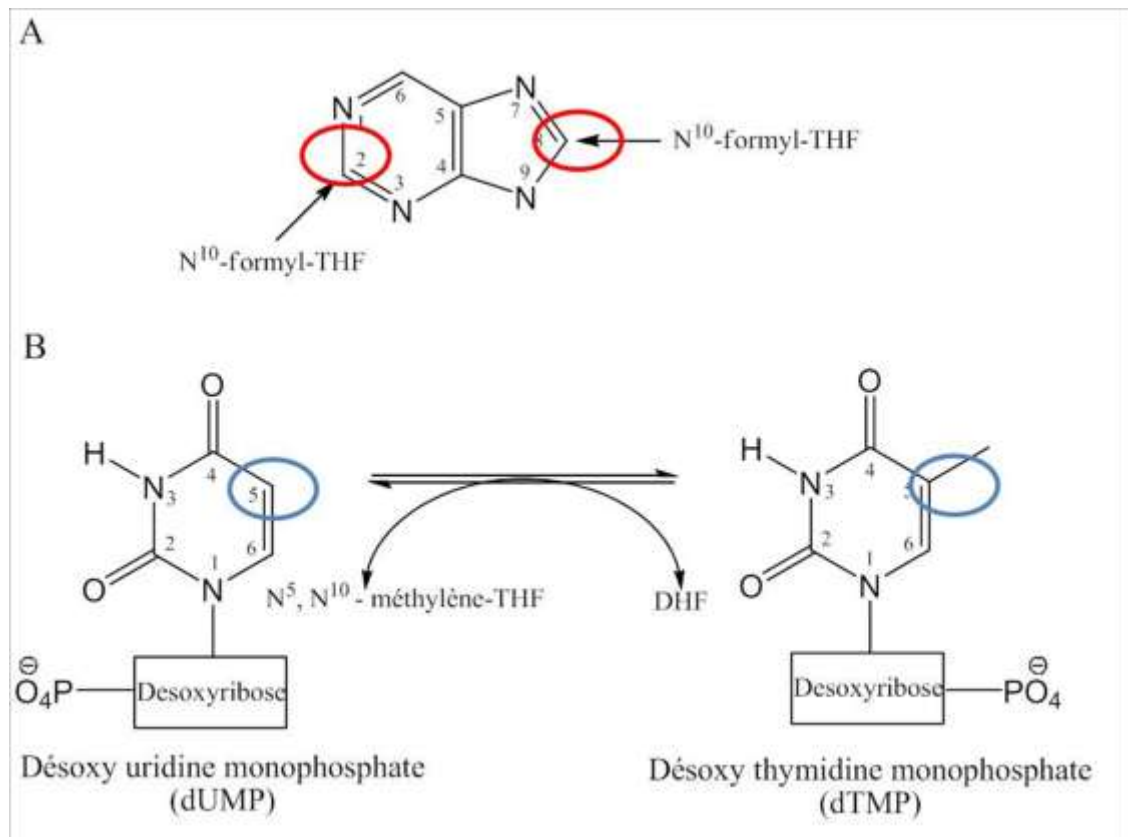
## Section 1.1.2 - Les dihydrofolates réductases

Les dihydrofolates réductases (DHFR) sont des enzymes vitales pour toutes les cellules. Chez les eucaryotes, les DHFR catalysent la réduction de l'acide folique en dihydrofolate (DHF) et subséquemment en tétrahydrofolate (THF) par l'oxydation du cofacteur nicotinamide adénine dinucléotide phosphate (NADPH) (figure 1.2). Contrairement aux humains<sup>22</sup>, les procaryotes peuvent synthétiser le DHF à partir du guanosine triphosphate et n'ont pas besoin de folate pour survivre. Par contre, la réduction de DHF en THF par la DHFR (encerclé en orange à la figure 1.2) est un goulot d'étranglement métabolique qui, si bloqué par l'action d'antifolates, ne peut être compensé par aucun processus métabolique. C'est ce qui explique l'efficacité des antifolates en tant qu'agents antiprolifératifs.



**Figure 1.2 : Structure du DHF, NADPH et de la réaction catalysée par la dihydrofolate réductase.** La liaison réduite lors de la réduction du DHF en THF et le carbone donneur d'hydrure sont encerclés en rouge et en bleu respectivement. L'étape de la réduction du folate en DHF n'est pas essentielle chez les bactéries. L'étape de réduction du DHF en THF, encerclé en orange, est essentielle chez les procaryotes et eucaryotes.

Le THF et ses dérivés sont en fait des donneurs de carbones et possèdent plusieurs rôles dans la cellule. Le THF est essentiel à la biosynthèse d'acide aminés comme la sérine et la méthionine. Le THF participe aussi à la synthèse de précurseurs de l'ADN (figure 1.3). Le THF entre dans la synthèse des purines (figure 1.3A), il fournit les carbones à la position 2 et 8 (encerclé en rouge). Le THF entre aussi dans la biosynthèse des désoxy thymidines monophosphates en fournissant le carbone qui forme le méthyl à la position 5 de la pyrimidine (figure 1.3B).



**Figure 1.3: Implication du THF dans la biosynthèse des purines et des désoxy thymidines monophosphates.** A) Les carbones de la position 2 et 8 (encerclé en rouge) sont fournis par le THF qui porte un groupement formyl sur son azote à la position 10. B) Le carbone du méthyl à la position 5 (encerclé en bleu) provient d'un méthylène situé les azotes 5 et 10 du THF.

### Section 1.1.3 - Résistance aux inhibiteurs de DHFR

Plusieurs mécanismes de résistance aux inhibiteurs de DHFR ont été identifiés jusqu'à maintenant chez les humains et les bactéries. Chez les humains, les mécanismes comprennent la déficience des transporteurs d'influx, l'augmentation de l'expression des transporteurs d'efflux, la surexpression de la DHFR et l'expression de DHFR mutantes.

Le folate ainsi que le MTX et le PTX entrent dans les cellules grâce aux transporteurs "*reduced folate carriers*" (RFC) (SLC19A1)<sup>23</sup>. Ce sont des transporteurs d'influx, c'est-à-dire qu'ils transportent leurs substrats du milieu extracellulaire vers le milieu intracellulaire. Une déficience chez ces récepteurs protège les cellules contre le MTX et PTX en les empêchant d'y entrer. Pour réduire l'activité des transporteurs RFC, plusieurs mécanismes moléculaires ont été identifiés. Les cellules peuvent exprimer des RFC mutants qui ont une moins grande affinité pour ces substrats<sup>24</sup>; la répression du gène du RFC dû à la perte de la fonction d'un ou plusieurs facteurs de transcriptions<sup>25</sup>; la méthylation du promoteur<sup>26</sup>; ou une altération dans le nombre de copies du gène RFC<sup>27</sup>. En contraste avec les transporteurs RFC, les transporteurs d'efflux "*multi- drug resistant*" MDR (MRP1 à MRP5) peuvent faire sortir le MTX ou le PTX de la cellule. Ainsi, une surexpression de ces transporteurs peut conférer une résistance à la cellule<sup>28</sup>.

Une surexpression de la DHFR peut aussi conduire à une résistance au MTX et PTX. C'est d'ailleurs un des mécanismes les plus fréquents<sup>29</sup>. Lorsque la quantité de molécules de DHFR augmente dans la cellule, la quantité non liée au MTX (ou PTX) augmente et ainsi une certaine quantité de DHFR peut produire le THF, permettant la survie de la cellule. Par ailleurs, il fut identifié, dans des lignées de cellules cancéreuses résistantes au MTX, des DHFR mutantes qui possédaient une moins grande affinité pour le MTX<sup>30</sup>. Malheureusement, une plus grande quantité de MTX ne peut être utilisée pour contrer ces mécanismes, dû à la toxicité du MTX en fortes concentrations.

Au niveau des bactéries, les mécanismes de résistance aux TMP identifiés sont causés par une imperméabilité de la paroi cellulaire, une surexpression de la DHFR

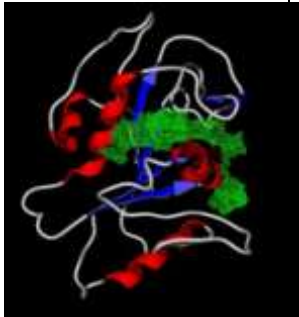
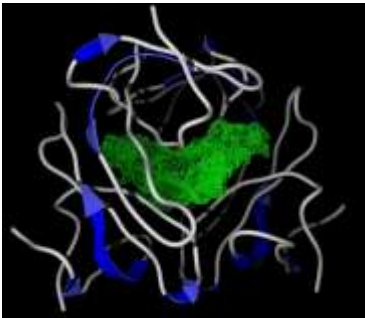
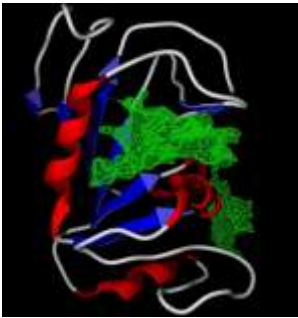
chromosomale, une expression de DHFR chromosomale mutante résistante au TMP ou encore une production de DHFR alternative, plasmidique et résistante au TMP, qui sera le sujet de cette étude.

Il a été démontré que certaines souches de bactéries comme *Pseudomonas aeruginosa* étaient intrinsèquement imperméables au TMP<sup>31</sup>. Il a été aussi reporté que la résistance au TMP basée sur l'imperméabilité de la paroi pouvait être transmise pour des souches comme *Klebsiella pneumoniae*<sup>32</sup>. D'autre part, tout comme pour la résistance au MTX des cellules eucaryotes, les bactéries peuvent surexprimer leur DHFR. Des mutations localisées au niveau du promoteur du gène peuvent augmenter le niveau de DHFR de 10 à 20 fois la normale dans les souches résistantes au TMP<sup>33</sup>. Une souche de *E. coli* ayant un taux 200 fois plus élevé de DHFR a déjà été identifiée<sup>34</sup>. Un mécanisme fréquent de résistance est l'expression de DHFR mutante ayant une affinité moins grande pour le TMP. Ces DHFR microbiennes chromosomales mutantes ont une concentration inhibitrice (IC<sub>50</sub>) envers le TMP de 46 à 300  $\mu\text{M}$ <sup>35</sup> en contraste avec les DHFR microbiennes chromosomales de type sauvage qui ont un IC<sub>50</sub> allant de 7 nM<sup>36</sup> à 400 nM<sup>37</sup> (il s'agit d'une comparaison approximative car la valeur de IC<sub>50</sub> peut différer selon les conditions expérimentales).

Les bactéries peuvent aussi produire une DHFR plasmidique hautement résistante au TMP : ce sont les DHFR de type II. Ces DHFR ont un K<sub>i</sub> pour le TMP plus d'un million de fois plus élevé que celui pour les DHFR microbiennes chromosomales de type I sauvage<sup>43, 44</sup>. Comme détaillé au tableau I, le K<sub>i</sub> pour le TMP des DHFR de type II, dont la DHFR R67 est la variante la mieux étudiée, est estimé à environ 150 $\mu\text{M}$ <sup>44</sup>, tandis celui de la DHFR microbienne chromosomale est de 20 pM<sup>43</sup>. Les DHFR de type II ont une structure totalement différente des DHFR de type I, même si elles catalysent la même réaction. La DHFR de type II est un homotétramère tandis que la DHFR de type I n'est constituée que d'un seul monomère, de plus, la structure des monomères n'a aucune relation évolutive. La grande résistance au TMP des DHFR de type II et leur nature plasmidique, qui permet le transfert horizontal du gène, rend cette enzyme particulièrement inquiétante. Des études épidémiologiques ont été entreprises

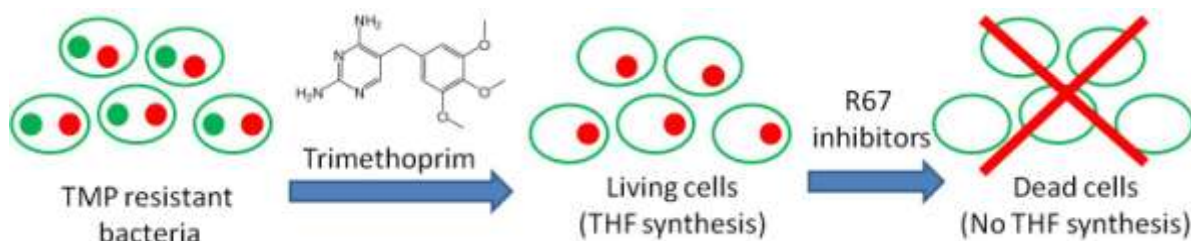
mais généralement les sondes utilisées ne sont pas spécifiques aux DHFR de type II<sup>38</sup>. On ne connaît donc pas l'étendue mondiale de ces enzymes.

**Tableau I : Caractéristique de la DHFR microbienne chromosomales, de la DHFR R67 et de la DHFR humaine**

	DHFR microbienne chromosomale	DHFR R67	DHFR humaine
Structure			
Origine	Bactérie- Chromosome	Bactérie - Plasmide	Humaine - Chromosome
Type	I	II	I
Masse	Monomère, 18 000Da <sup>16</sup>	Tétramère, 34 000 Da <sup>16</sup>	Monomère, 21 453 Da <sup>39</sup>
$K_M^{DHF}$	1.2 $\mu M$ <sup>40</sup>	8.2 $\mu M$ <sup>41</sup>	0.12 $\mu M$ <sup>42</sup>
$K_M^{NADPH}$	0.94 $\mu M$ <sup>40</sup>	1.6 $\mu M$ <sup>41</sup>	0.16 $\mu M$ <sup>42</sup>
$K_i^{TMP}$	20 pM <sup>43</sup>	150 $\mu M$ <sup>44</sup>	5.2 $\mu M$ <sup>42</sup>

Comme il a été mentionné, une des DHFR de type II les plus étudiées est la DHFR R67 (voir tableau I). Comme pour les autres DHFR de type II, c'est un

homotétramère dont la structure n'est pas apparentée à celle des DHFR de type I et elle a une affinité très faible pour le TMP. Ainsi, lorsque la DHFR chromosomale bactérienne est inhibée par le TMP, la DHFR R67 produit le THF afin d'assurer la survie de la bactérie même en présence de TMP. Ainsi, des inhibiteurs de DHFR R67 administrés avec le TMP pourraient contrer ce mécanisme et empêcher la survie des bactéries, qui expriment la DHFR R67, en présence de TMP (voir la figure 1.4).

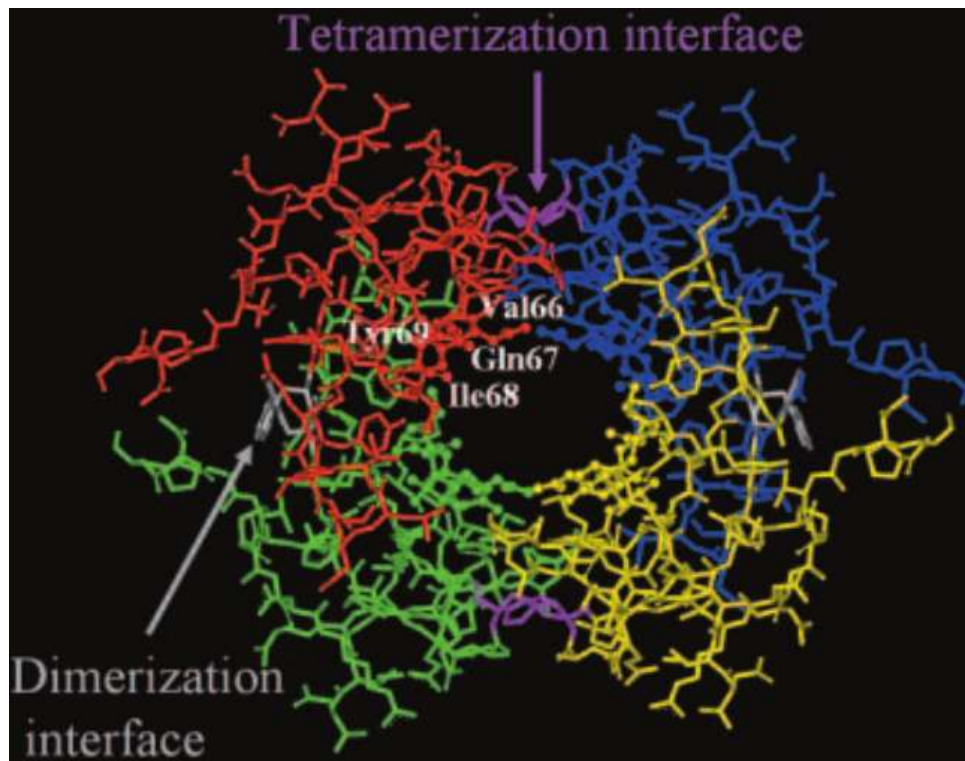


**Figure 1.4 : Concept schématisé d'inhibiteurs de DHFR R67 pour contrer la résistance au TMP conféré par la DHFR R67.** Les cercles pleins verts représentent la DHFR chromosomale et les cercles rouges représentent la DHFR R67. Le TMP inhibe la DHFR chromosomale : la DHFR R67 n'étant pas affectée par le TMP, elle permet la production de THF et la survie de la bactérie. L'ajout d'inhibiteur de la DHFR R67 inhiberait cette dernière. La cellule perdrait son moyen de produire du THF et mourrait.

## Section 1.2 - La dihydrofolate réductase R67

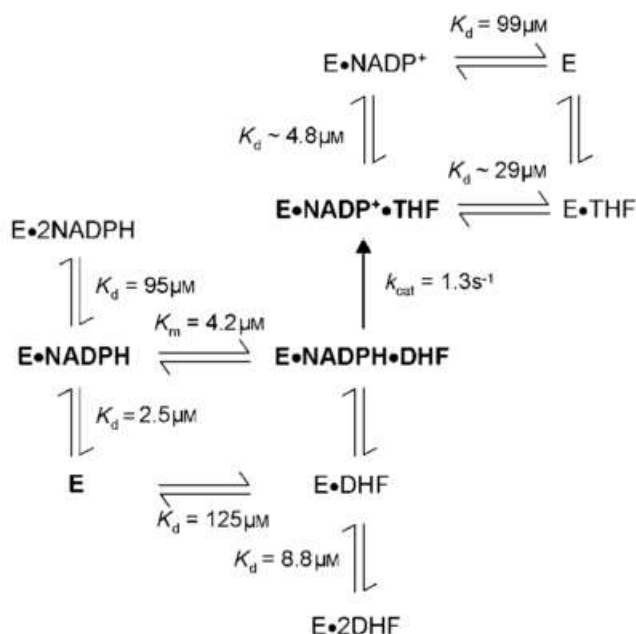
Bien que les DHFR de type I et de type II catalysent la même réaction, elles n'ont pas la même efficacité. Les DHFR de type II catalysent moins efficacement la réduction du DHF en THF. L'étape limitante dans la réaction catalysée par la DHFR chromosomale est le relâchement du THF avec une constante catalytique ( $k_{cat}$ ) de  $29 \text{ s}^{-1}$  tandis que pour la DHFR R67, l'étape limitante est le transfert de l'hydrure du NADPH vers le DHF avec un  $k_{cat}$  de  $1.3 \text{ s}^{-1}$ . Une autre différence réside dans le fait que les DHFR II, dont la DHFR R67, sont des homotétramères. La figure 1.5<sup>41</sup> illustre la structure quaternaire de la DHFR R67.





**Figure 1.5 : Structure quaternaire de la DHFR R67.** Les monomères sont colorés en bleu, jaune, vert et rouge respectivement. L'interface de dimérisation implique des contacts hydrophobes. L'interface de tétramérisation implique les résidus histidine 62. Les résidus Val66, Gln67, Ile68 et Tyr69 sont les résidus du monomère qui forment le site actif. La figure est tirée de la référence 42.

Le monomère est formé de 5 feuilletts bêta antiparallèles. Deux monomères se lient ensemble pour former un dimère par des contacts hydrophobes. Ensuite, deux dimères se lient ensemble pour former un tétramère par les résidus His62. Cette tétramérisation forme un tunnel au centre de l'enzyme : c'est le site actif de la DHFR R67. Le site actif de la DHFR R67 est principalement formé des résidus Val66, Gln67, Ile68 et Tyr69 et il est généralement hydrophobe. Une des particularités de ce site actif est qu'il possède une symétrie 222 et qu'il est juste assez grand pour accepter le substrat et son cofacteur. Donc, peu importe le côté du tunnel que le substrat et cofacteur vont utiliser pour entrer dans le site actif, ils formeront les mêmes interactions avec la DHFR R67. Si le NADPH entre du côté gauche, le DHF entrera du côté droit et l'inverse est aussi vrai. Ainsi, le DHF et le NADPH partagent le même site de liaison. En effet, la DHFR R67 peut lier deux molécules de DHF, deux molécules de NADPH ou une molécule de DHF et une de NADPH (voir figure 1.6<sup>45</sup>). Seule cette dernière donne lieu à la liaison productive.

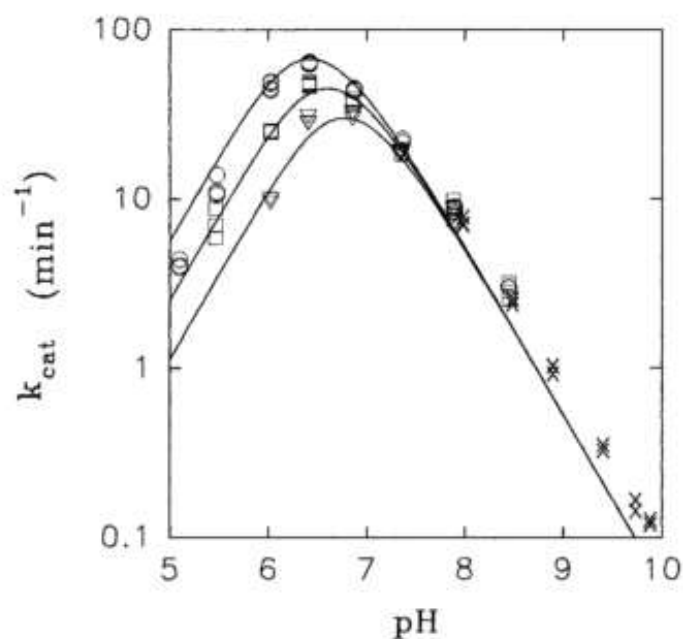


**Figure 1.6: Mécanisme de catalyse proposé pour la DHFR R67.** La voie favorisée est montrée en caractères gras. La figure est tirée de la référence 48.

La DHFR R67 libre peut se lier soit au DHF ou au NADPH. La liaison du NADPH est favorisée, la constante de dissociation pour le complexe DHFR R67•NADPH étant environ 50 fois plus petite que pour le complexe DHFR R67•DHF. Ensuite, le complexe DHFR R67•NADPH peut se lier avec une molécule de NADPH ou de DHF. Il existe une coopérativité et la liaison avec une molécule de DHF est favorisée par rapport à celle d'une deuxième molécule de NADPH. La molécule de DHF doit être protonnée préalablement par le milieu réactionnel, car la DHFR R67 ne possède pas de donneur de proton<sup>46</sup>. Le complexe DHFR R67•NADPH•DHF est formé et l'hydrure est transféré du NADPH au DHF pour ainsi former le THF. Les deux autres possibilités de complexes, DHFR R67•2NADPH et DHFR R67•2DHF, ne sont pas des complexes actifs.

La structure particulière de la DHFR R67 ainsi que son mécanisme de catalyse qui dépend du pH du milieu réactionnel, donnent à la DHFR R67 un profil de pH assez particulier. En effet, elle est très sensible au pH, comme on peut le voir à la figure 1.7<sup>47</sup>. L'activité maximale de la DHFR R67 se trouve environ au pH 6.5. Pour une variation de

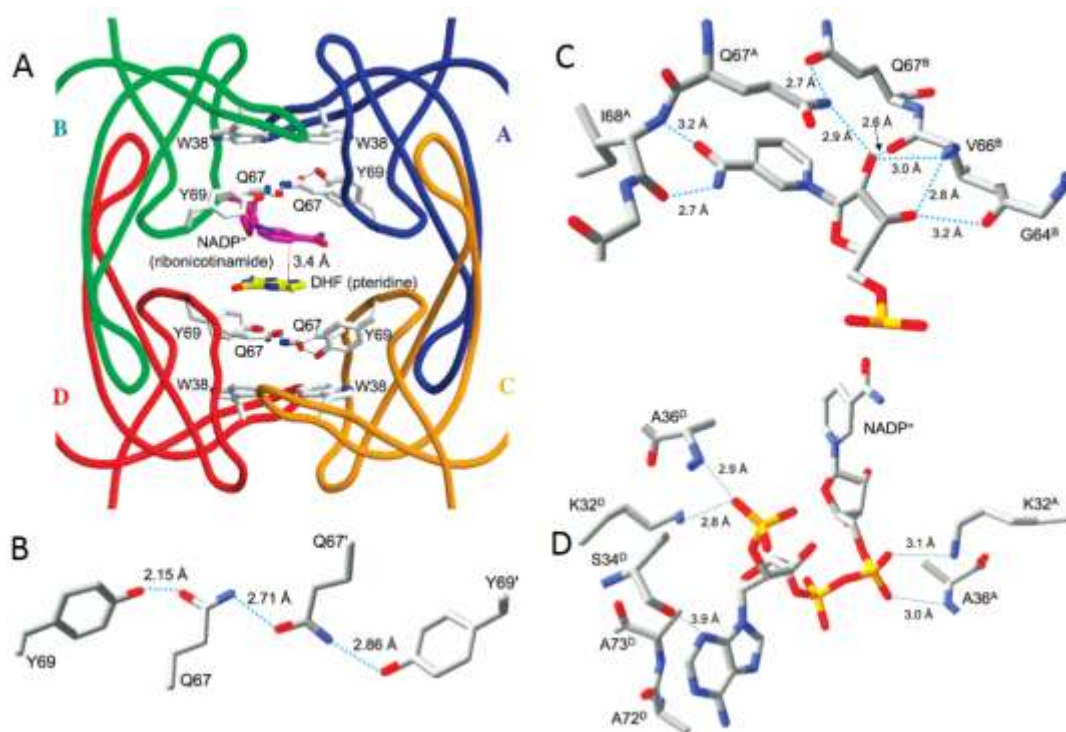
$\pm 1$  unité de pH, le  $k_{cat}$  de la DHFR R67 passe de  $90 \text{ min}^{-1}$  à environ  $10 \text{ min}^{-1}$ . La chute de l'activité à pH basique provient du fait que la DHFR R67 ne possède pas de donneur de proton et que le DHF doit être protonné par le milieu réactionnel. Ainsi, plus le pH est acide, plus le DHF sera protonné et plus l'activité augmente. Par contre, les résidus His62 situés au niveau de l'interface de tétramérisation seront majoritairement protonnés vers un pH de 6.5 ou moins, ce qui provoquera la dissociation du tétramère, donc une chute de l'activité.



**Figure 1.7: Influence du pH sur l'activité de la DHFR R67.** Les cercles, carrés et triangles représentent une concentration de DHFR R67 de 828 nM, 370 nM et 166 nM respectivement. La figure est tirée de la référence 50.

## Section 1.2.1 - Site actif de la DHFR R67

À la section précédente, il a été mentionné que le site actif de la DHFR R67 était principalement constitué des résidus Val66, Gln67, Ile68 et Tyr 69 provenant de chacun des quatre monomères constituant l'homotétramère. À la figure 1.8<sup>48</sup>, on peut voir le rôle que ces résidus jouent dans la liaison du DHF et du NADPH au site actif de la DHFR R67. Ce site actif est inhabituel : alors qu'un homotétramère comporterait habituellement quatre sites actifs équivalents, celui-ci ne possède qu'un seul site actif central. Ceci a pour conséquence qu'un même résidu - par exemple, la Gln67 - provenant des quatre monomères jouera quatre rôles différents lors d'un même événement catalytique.



**Figure 1.8 : Implication des résidus clés dans la liaison du DHF et du NADPH dans le site actif de la DHFR R67.** A) Motif «sandwich» observé dans le site actif de la DHFR R67 lors de la liaison avec le DHF et NADPH. B) Réseau de ponts H entre les résidus Gln67 et Y69 entre deux monomères qui forme une "pince" moléculaire. C) Interaction impliquant la partie ribonicotinamide du NADPH et la DHFR R67. D) Interactions impliquant la partie adénosine diphosphate du NADPH et la DHFR R67 à l'entrée du site actif. La figure est tirée de la référence 51.

La structure tertiaire de la DHFR R67 est formée par une organisation de plusieurs couches d'interactions d'empilement, ce qui forme un motif « sandwich »<sup>48</sup> (figure 1.8A). Sur la couche extérieure du motif, les Trp38, via leurs noyaux indole, forment des interactions d'empilement avec les cycles aromatiques des Tyr69, ce qui permet de former un réseau de ponts H (figure 1.8B) entre les alcools des Tyr69 et les carbonyles des chaînes latérales des Gln67. Ce réseau de ponts H agit comme une pince moléculaire qui tient en place le groupement ptérine du DHF et le groupement nicotinamide du NADPH au centre du site actif de la DHFR R67. Des interactions d'empilement sont aussi présentes entre les groupements ptérine du DHF et nicotinamide du NADPH. L'empilement entre le groupement ptérines et le nicotinamide est favorisé par rapport à l'empilement entre deux ptérine ou deux nicotinamides, ce qui explique la coopérativité qui favorise la liaison du DHF au lieu d'une deuxième molécule de NADPH<sup>45</sup>.

Aux figures 1.8C et 1.8D, les atomes impliqués dans les interactions entre le NADPH et la DHFR R67 sont visibles. En 1.8C, on voit les interactions que le ribonicotinamide forme avec le site actif de la DHFR R67. La chaîne principale de la Val66 (NH) forme des ponts H avec les groupements alcool en position 2' et 3' (voir figure 1.2 pour la structure du NADPH et DHF) du ribose du ribonicotinamide. La chaîne latérale de la Gln67 (CO) forme des ponts H avec l'alcool en position 2' du ribose du ribonicotinamide. La chaîne principale de l'Ile68 (NH et CO) reconnaît le motif amide retrouvé sur le nicotinamide du NADPH et forme des ponts H avec ce groupement (NH de la chaîne principale avec CO du nicotinamide et CO de la chaîne principale avec NH<sub>2</sub> du nicotinamide). En 1.8D, on peut voir les interactions que la DHFR R67 forme avec la partie adénosine diphosphate de NADPH. Les NH<sub>3</sub><sup>+</sup> situé sur les Ly32 qui sont à l'entrée du site actif, forment des interactions électrostatiques avec les phosphates chargés négativement. Les chaînes principales des Ala36 (NH) forment des liaisons H avec les oxygènes des groupements phosphates.

Au contraire du NADPH, dont l'entière structure a été résolue, seul le groupement ptérine du DHF a été résolu dans la structure tertiaire de la DHFR R67.

Ainsi, il n'est pas possible de décrire précisément les interactions que le DHF forme avec la DHFR R67. Par contre, le NADPH et le DHF ont le même site de liaison, donc il est attendu que les interactions et les résidus impliqués soient similaires. En effet, des études de mutagénèses ont montré que les résidus Lys32 présents à l'entrée du site actif forment des interactions électrostatiques avec le glutamate du DHF<sup>49</sup> de façon similaire à l'interaction entre les Lys32 et les phosphate du NADPH. La chaîne principale des Ile68 reconnaît aussi le motif amide présent sur le groupement pterine du DHF et forme des ponts H avec ce dernier (NH de la chaîne principale avec CO du pterin et CO de la chaîne principale avec NH du pterin)<sup>50</sup>.

La perte d'affinité du TMP pour la DHFR R67 s'explique en partie par le fait que le TMP ne possède pas d'amide; ainsi, le TMP ne forme pas d'interaction avec la chaîne latérale de l'Ile68<sup>51</sup>. Une autre raison qui explique l'affinité plus faible du TMP pour la DHFR R67 est l'absence d'un donneur de proton dans son site actif (glutamate ou aspartate). Lors de la liaison du TMP au site actif de la DHFR microbienne chromosomale ( $K_i = 20 \text{ pM}$ ), le TMP est chargé positivement et forme une liaison électrostatique avec le carboxylate d'un résidu glutamate situé dans son site actif, ce qui n'est pas possible avec la DHFR R67<sup>48</sup>.

## Section 1.2.2 - Inhibiteurs connus de la DHFR R67

Jusqu'à ce jour, deux inhibiteurs de la DHFR R67 sont connus : le rouge de congo et la novobiocin (figure 1.9). Le  $K_i$  du rouge de congo est de  $2\mu\text{M}^{16}$  et celui de la novobiocin est de  $70\mu\text{M}^{16}$ . Les caractéristiques structurales qui peuvent expliquer la différence de puissance des deux inhibiteurs seront discutées ci-dessous.

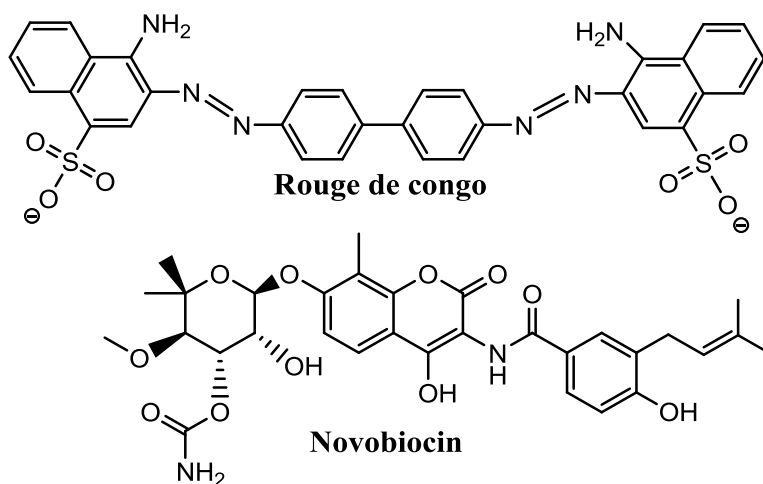


Figure 1.9 : Inhibiteurs connus de la DHFR R67

Dans les deux cas, les structures sont plutôt linéaires et possèdent plusieurs groupements aromatiques ainsi les molécules possèdent beaucoup de carbones hybridés  $sp^2$ . La présence de cycles aromatiques pourrait permettre la formation d'empilements intermoléculaires qui sont semblables aux interactions importantes dans la liaison du DHF et du NADPH à la DHFR R67 (section 1.2.1). Le rouge de congo possède aussi une partie centrale plus hydrophobe dû au fait qu'il possède moins d'hétéroatomes que la novobiocin. Une structure plus hydrophobe pourrait être avantageuse, car le site actif de la DHFR R67 est relativement hydrophobe. Le rouge de congo possède des groupements chargés négativement, contrairement à la novobiocin. Des groupements chargés négativement pourraient être avantageux car ils pourraient reproduire les interactions entre le glutamate du DHF et les phosphates du NADPH et les Lys32. Finalement, contrairement à la novobiocin, le rouge de congo est symétrique, ce qui peut être un avantage vu la symétrie du site actif de la DHFR R67. Ainsi, les

caractéristiques qui semblent avantager la liaison d'un ligand au site actif de la DHFR R67 sont les suivantes : l'hydrophobie, la linéarité, la planarité, la symétrie, et le fait de posséder des cycles aromatiques et des charges négatives.

La novobiocin est un inhibiteur de gyrase d'ADN bactérienne. Cette dernière utilise l'énergie de l'ATP pour catalyser le surenroulement de l'ADN. Cette fonction est essentielle pour la transcription et la réplication de l'ADN et est assumée par des enzymes différentes chez les eucaryotes ce qui fait de la gyrase bactérienne une cible intéressante. Dû à sa faible puissance, sa faible biodisponibilité orale<sup>52</sup> et sa forte masse moléculaire (612 g/mol), la novobiocin n'a pas servi d'inspiration pour le développement d'inhibiteurs pour la DHFR R67.

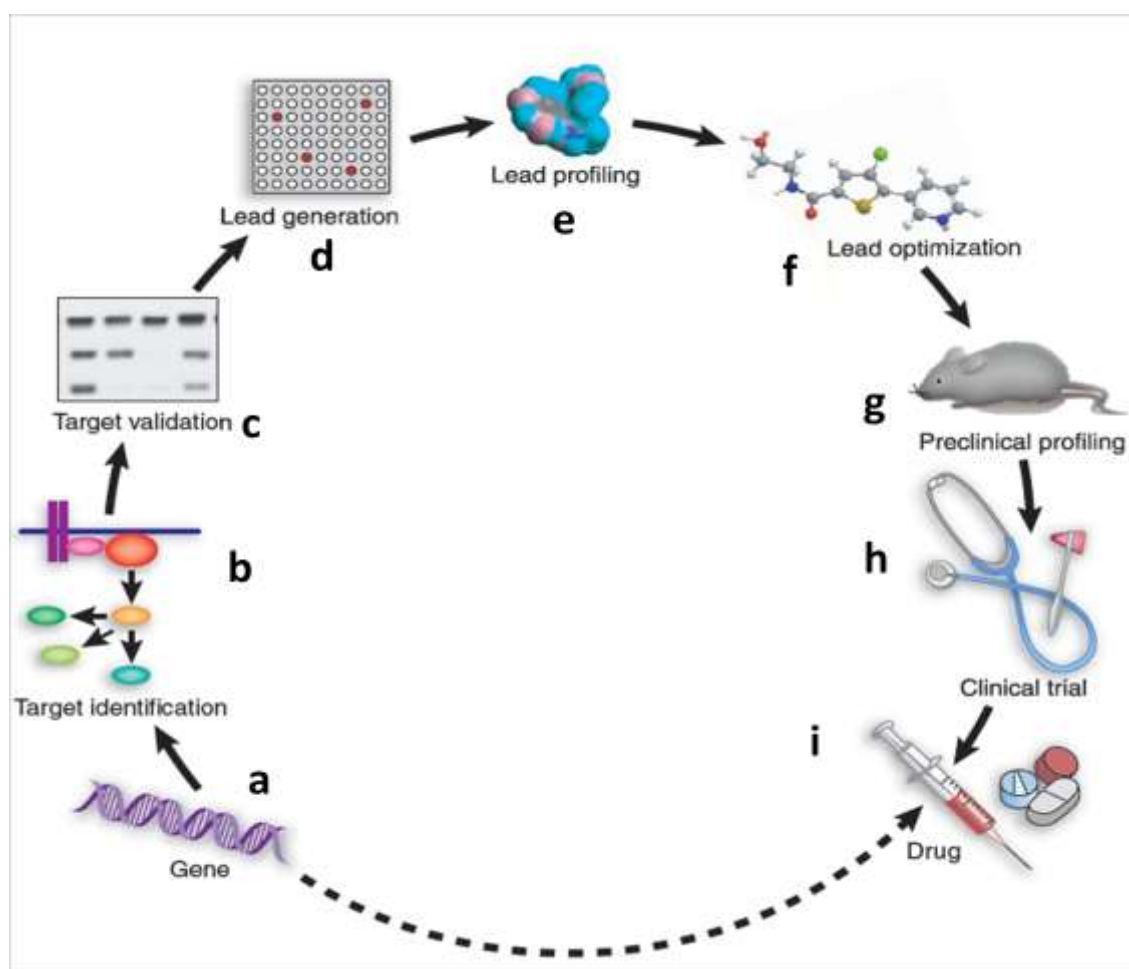
Le rouge de congo est une teinture fréquemment utilisée pour colorer des coupes en microscopie, ce qui démontre que la molécule peut interagir avec toutes sortes d'autres cibles. Sa toxicité *in vivo* a été démontrée et Santé Canada considère cette substance comme étant toxique<sup>53</sup>. Plus précisément, ce composé est connu pour produire des problèmes au niveau de la fertilité et du développement humain. Il est connu aussi pour avoir des effets carcinogènes et des effets mutagènes<sup>54</sup>. Bien que le rouge de congo démontre une puissance intéressante envers la DHFR R67, il est hautement toxique. Le développement d'un inhibiteur pour la DHFR R67 à partir du rouge de congo n'a donc pas été entrepris et le développement *de novo* d'inhibiteurs de la DHFR R67 a été amorcé. Dans la prochaine section, nous ferons un survol du processus de développement du médicament pour comprendre la logique et situer les travaux qui ont été effectués.



## Section 1.3 - Du gène au médicament

Les étapes du développement du médicament sont représentées à la figure 1.10<sup>55</sup>. Ces étapes seront discutées en détail dans les sections subséquentes. La première phase (figure 1.10a) consiste à comprendre les mécanismes moléculaires d'une maladie. À la deuxième (figure 1.10b), il faut identifier une stratégie pour contrer cette maladie. Cela consiste en l'identification de la cible. L'étape 3 (figure 1.10c) consiste à valider la cible choisie en évaluant son efficacité pour le traitement recherché. À la quatrième étape (Figure 1.10d), on cherche des molécules qui, interagissant avec la cible, pourront provoquer l'effet recherché (inhibition ou activation). Durant cette étape, il sera généralement possible d'identifier plusieurs classes de molécules. À l'étape 5 (figure 1.10e), la classe de composés la plus prometteuse est déterminée en évaluant divers facteurs tels que la puissance, la sélectivité, le mécanisme d'action, l'absence de groupements chimiques indésirables et la facilité de synthèse d'analogues. Pour poursuivre les efforts de développement, on passe à la phase d'optimisation (figure 1.10f), où l'on améliore les propriétés déterminées lors de l'étape de caractérisation. On amorce alors des études de pharmacologie *in vitro* et *in vivo* pour l'évaluation de paramètres pharmacocinétiques comme la clairance, la perméabilité, la biodisponibilité, la demi-vie, la toxicité et l'effet sur les cytochromes P450 qui sont des enzymes responsables de la dégradation des molécules exogènes. Quand des molécules ayant des paramètres satisfaisants ont été identifiées, on passe aux essais précliniques (figure 1.10g). C'est lors de ces essais entrepris à l'aide d'animaux modèles, qu'on détermine pour la première fois l'efficacité thérapeutique des molécules : il s'agit de la preuve de concept. Durant ces essais précliniques, les doses thérapeutique et létale sont aussi déterminées ainsi qu'un profil pharmacocinétique complet en vue des tests sur les humains. Enfin, une molécule sera choisie comme candidat-médicament ("*drug candidate*") afin d'être testée chez l'humain lors des études cliniques. Celles-ci (figure 1.10h) commencent à l'aide d'un nombre restreint d'humains en bonne santé afin de mesurer les paramètres pharmacocinétiques chez l'humain ainsi que l'apparition d'effets secondaires importants. Par la suite, le candidat-médicament est testé chez un nombre restreint de patients atteints de la maladie d'intérêt afin de mesurer son efficacité thérapeutique de même que pour déterminer la dose thérapeutique chez l'humain. Le

candidat-médicament sera testé sur des groupes de plus en plus nombreux afin de rendre les tests statistiques plus fiables dans le but d'identifier des effets secondaires peu fréquents s'il y a lieu. Finalement, on peut procéder à la demande d'une autorisation de mise sur le marché. Lorsqu'elle aboutit, le candidat-médicament alors devenu un médicament sera vendu (figure 1.10i).

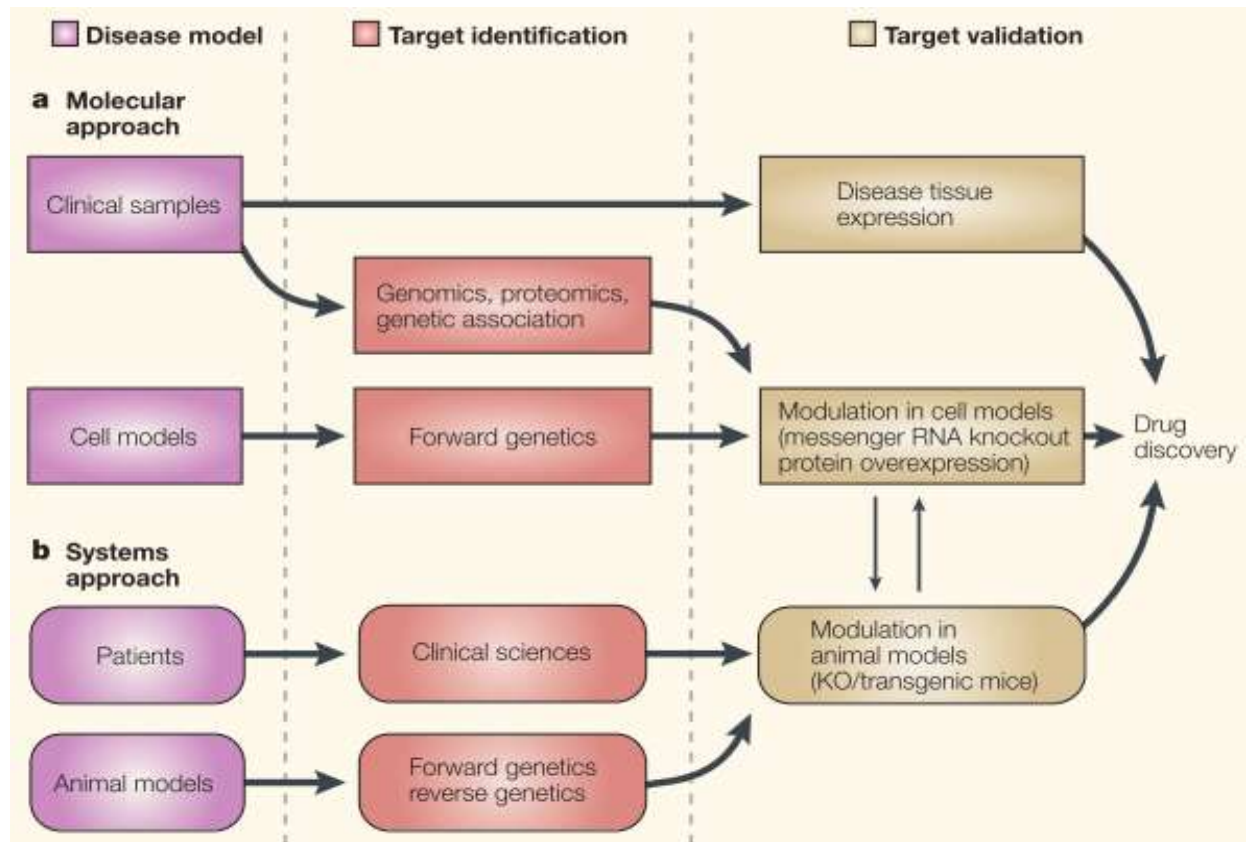


**Figure 1. 10: Vue d'ensemble du processus du développement d'un médicament.** Voir le texte et les sous-sections pour les détails de chaque étape. La figure est tirée de la référence 57.

### Section 1.3.1 - Découverte de la cible

Cette section présentera les étapes de la découverte d'une cible , dont le choix des tissus/animaux modèles (figure 1.10a), l'identification de la cible (figure 1.10b) et la validation de la cible (figure 1.10c). La figure 1.11<sup>56</sup> représente les deux approches

utilisées dans la découverte de nouvelles cibles. Il s'agit des approches moléculaire et physiologique. La différence principale entre ces approches résulte notamment du choix des tissus/animaux modèles qui dictent le reste du processus.



**Figure 1.11 : Vue d'ensemble du processus de découverte d'une nouvelle cible.** Deux approches sont présentées : l'approche moléculaire A) et l'approche physiologique B) La figure est tirée de la référence 58.

L'approche moléculaire se fait au niveau microscopique et est axée sur des études *in vitro* de cellules provenant de patients (échantillon clinique) ou de lignées cellulaires qui simulent la pathologie d'intérêt. Ces dernières offrent la possibilité d'étudier les mécanismes cellulaires impliqués dans la pathologie d'intérêt. Toutefois les cellules étudiées *in vitro* ne se retrouvent pas dans un environnement naturel et par conséquent, il est difficile de prédire si le comportement d'une cible moléculaire demeurera le même au niveau de l'organisme<sup>57</sup>. Dans l'approche physiologique, l'étude

de la pathologie d'intérêt se fait au niveau macroscopique en étudiant l'organisme soit de patients atteints de la maladie d'intérêt ou d'animaux modèles chez lesquels la maladie est reproduite. L'avantage de cette approche est la possibilité d'étudier des maladies observables uniquement au niveau de l'organisme, par exemple l'obésité, l'hypertension ou les maladies neurodégénératives. Idéalement, étudier des patients malades et sains serait la meilleure approche dû au fait qu'on étudierait la maladie directement dans l'organisme humain. Par contre, dû à des raisons d'éthique et de coûts, on reproduit généralement la maladie d'intérêt dans un animal modèle. Ceci peut être un désavantage car les causes d'une maladie peuvent différer entre les animaux modèles et l'humain<sup>56</sup>. De plus, certaines maladies sont difficiles ou même impossibles à étudier avec des animaux modèles, telles que les maladies d'ordre psychologique.

Lors de l'étape d'identification, on tente généralement de reconnaître une ou des protéines qui modulent la progression de la maladie d'intérêt. Il s'agit d'observer une corrélation entre le niveau d'expression ou la présence d'une mutation d'une protéine dans les cellules provenant des échantillons cliniques ou des modèles étudiés. Dans l'approche moléculaire, des techniques de génomique<sup>58</sup>, protéomique<sup>59</sup> et d'association génétique<sup>60</sup> sont utilisées. En génomique, on tente de déterminer si un gène est surexprimé ou sous-exprimé en quantifiant l'ARNm des différents gènes. Il est ainsi possible de cribler tous les gènes en une seule expérience dû à la réussite du Projet Génome Humain<sup>56</sup>. Toutefois on observe généralement des centaines de gènes ayant des taux d'expression différents des contrôles, ce qui mène à la détermination de nombreux faux positifs. De plus, un taux d'ARNm différent ne mène pas toujours à une expression différente des protéines<sup>56</sup>. Dans le même ordre d'idées, en protéomique, on mesure directement le niveau d'expression des protéines dans un échantillon<sup>59</sup>. Toutefois, l'absence d'étape d'amplification (comme la PCR pour la génomique) rend difficile la détection des protéines faiblement exprimées<sup>61</sup>. Finalement, l'association génétique, contrairement aux techniques génomiques et protéomiques, permet l'identification de variation de séquence chez une protéine dans les échantillons, dans le cas où la cible est préalablement connue.

Sur les lignées cellulaires modèles et sur les animaux modèles, il est possible d'effectuer des études de génétique classique ("*forward genetics*") qui consiste à aller du phénotype vers le génotype. Ainsi, il est possible de cribler de grandes bibliothèques chimiques à la recherche de molécules qui peuvent influencer le phénotype de la pathologie, donnant lieu à l'identification de molécules actives contre la maladie d'intérêt<sup>62</sup> et éventuellement d'une cible. Finalement, sur les animaux modèles, la génétique inverse, qui consiste à introduire des mutations ou le knockout de gènes, peut être mise en pratique pour identifier les gènes qui influencent le phénotype de la pathologie d'intérêt<sup>63</sup>.

La découverte des DHFR de type II en milieu clinique (dont la DHFR R67 fait partie) s'est faite par l'approche moléculaire. En 1972, Fleming rapportait la découverte de plasmides R qui conféraient la résistance au TMP<sup>64</sup> chez des bactéries de type *E.Coli* prélevées dans l'urine de patients infectés. Par la suite, le mécanisme moléculaire par lequel ces plasmides conféraient la résistance au TMP a été élucidé par Amyes et Smith<sup>65</sup> et Skold et Widh<sup>66</sup>. Ils ont démontré que ce mécanisme implique une seconde espèce de dihydrofolate réductase hautement résistante au TMP; ce sont les dihydrofolates de type II. Plus tard, en 1976, Amyes et Smith<sup>44</sup> ont isolé et caractérisé cette enzyme, ils ont démontré que cette enzyme avait un IC<sub>50</sub> pour le TMP environ 25 000 fois plus élevé que pour la DHFR chromosomale. De plus, ils ont déterminé que l'enzyme de type II avait une masse moléculaire de 35 000 Da, ce qui est nettement supérieure à la DHFR chromosomale d'*E.Coli* avec sa masse de 21 000 Da<sup>44</sup>. À partir de la masse moléculaire, ils ont émis une hypothèse sur l'origine des DHFR de type II. Dû à la grande différence de masse entre les deux DHFR, ils serait surprenant que les DHFR de type II proviennent d'*E.Coli*<sup>44</sup>. Les DHFR de type II pourrait provenir d'un bactériophage dû au fait que certains bactériophages, comme le T<sub>6</sub>, possèdent une DHFR avec un poids moléculaire de 31 000 Da<sup>44</sup>. Par contre, encore aujourd'hui, on ne connaît pas avec certitude l'origine des DHFR de type II.

Une protéine dont la fonction est déterminante dans une pathologie ne fait pas nécessairement une bonne cible. En effet, plusieurs autres facteurs doivent être pris en

compte<sup>67</sup>. Idéalement, la cible ne doit pas être exprimée dans tous les tissus du corps humain pour réduire les effets du médicament sur les tissus sains<sup>67</sup>. Dans le cas où la cible est exprimée dans plusieurs tissus, la modulation de cette dernière doit avoir un plus grand impact dans les tissus malades que ceux sains<sup>67</sup>. Si l'on considère l'aspect technique, il doit être possible de développer un essai permettant le criblage à haut débit de la cible, dans le but de découvrir des molécules actives<sup>67</sup>. La structure de la cible doit préférentiellement contenir des cavités qui pourront être exploitées pour permettre la liaison de molécules<sup>67</sup>.

En plus de son rôle clé dans la résistance au TMP, la DHFR R67 a été définie comme cible dû aux caractéristiques suivantes. Elle possède un site actif avec un grand volume libre qui peut être exploité pour la liaison à d'inhibiteurs (section 1.1.3). De plus, cette protéine est dotée d'une structure évolutivement distincte de la DHFR humaine, ce qui permettrait l'obtention d'inhibiteurs sélectifs à la DHFR R67 (section 1.2.1). Enfin, un essai compatible avec le criblage à haut débit a pu être développé pour la DHFR R67 (voir chapitre 2).

La validation de la cible permettra de démontrer le rôle de la cible identifiée et sa valeur thérapeutique. Il s'agit de moduler l'activité de la protéine cible dans le tissu sain d'intérêt afin de reproduire la pathologie dans un contexte *in vitro* et *in vivo* (ou l'inverse, moduler l'activité de la cible dans un tissu malade pour modifier le phénotype de la pathologie). Pour se faire, on module l'activité de la protéine cible en surexprimant soit un mutant négatif (diminution de l'activité), soit la protéine de type sauvage (augmentation de l'activité), ou encore en utilisant des réactifs qui interagissent avec la cible comme des anticorps<sup>68</sup>. Le désavantage de ces techniques réside dans le fait que la haute concentration de protéine ou la présence d'un réactif exogène en grande quantité peut influencer le fonctionnement des cellules. Le contrôle du niveau d'activité de la protéine d'intérêt peut aussi se faire au niveau de son ARNm en utilisant des ARN d'interférence (ARNi) qui bloquent la traduction de l'ARNm d'intérêt. Les ARNi sont très spécifiques et ne requièrent que les informations au niveau du gène d'intérêt et non au niveau de la structure de la cible<sup>56</sup>. Par contre, cette technique sert uniquement à réduire le niveau d'expression de la cible. Une souche de bactérie sensible au TMP,

de l'espèce *E. coli* qui n'exprime pas la DHFR R67, a une concentration minimale inhibitrice (CMI) de moins de 0.5  $\mu\text{g/mL}$ <sup>69</sup>. Une autre souche de *E. coli*, exprimant la DHFR R67, a une CMI plus grande que 1000  $\mu\text{g/mL}$ <sup>64</sup>. Cela démontre que l'inhibition de la DHFR R67 permettrait de rétablir la sensibilité des cellules au TMP. Ainsi, la DHFR R67 est une cible adéquate afin de briser la résistance au TMP conférée par cette enzyme.

### Section 1.3.2 - Identification de composés actifs

Lorsqu'une cible a été identifiée et validée, on passe à l'étape de recherche de composés actifs (figure 1.10d). Premièrement, on développe les essais qui permettront le test de composés. Deuxièmement, on doit choisir une approche de recherche de composés. Finalement, on teste les composés.

Plusieurs techniques peuvent être utilisées dans le but de mettre au point un essai, par exemple, la spectroscopie de masse (MS)<sup>70,71</sup>, la résonance magnétique nucléaire (RMN)<sup>72</sup>, la cristallographie par diffraction de rayons X<sup>73</sup>, la titration calorimétrique (ITC)<sup>74</sup>, la résonance de plasmon de surface<sup>75</sup> ou encore les essais d'activité.

La spectroscopie de masse peut servir à déterminer les interactions entre des petites molécules et des biomacromolécules telles que les protéines ou les ARN<sup>71</sup>. Cette technique consiste à observer la variation du ratio masse/charge ( $m/z$ ) résultat de sa liaison avec un composé. La masse du complexe cible/composé, permet d'identifier le composé. Ainsi, cette technique permet de tester simultanément plusieurs composés avec plusieurs cibles en un seul essai<sup>70</sup>. Toutefois cette technique n'est présentement pas fréquemment utilisée et elle est fiable seulement pour mesurer des interactions relativement fortes avec une limite de détection dans les bas micromolaires. De plus, elle nécessite beaucoup de matériel. Vu la nature tétramérique de la DHFR R67, cette technique ne semble pas appropriée du fait que la phase d'ionisation pourrait briser le tétramère et affecter la liaison des ligands.

La résonance magnétique nucléaire (RMN) est utilisée fréquemment dans la recherche de composés actifs. Cette technique consiste à observer les déplacements chimiques  $^{15}\text{N}$  et  $^1\text{H}$ , dans un spectre à deux dimensions HSQC ("*heteroatom single quantum correlation*") de la cible en présence et en absence de composés. Les variations de déplacements chimiques suggéreront un événement de liaison. Cette technique permet d'obtenir des informations structurales au niveau des résidus impliqués dans la liaison d'un composé. De plus, elle est suffisamment sensible pour déterminer de faibles affinités (ordre du millimolaire)<sup>72</sup>. Par contre cette méthode ne permet pas d'analyser des cibles de très haut poids moléculaire et il peut être fastidieux d'analyser les résultats. Enfin, elle requiert de hautes concentrations de protéines marquées  $^{15}\text{N}$  et de composés, ce qui peut engendrer des problèmes de solubilité et de coûts<sup>75</sup>. Il est aussi possible de tester plusieurs composés à la fois; par contre, il sera nécessaire de raffiner l'expérience afin d'identifier le ou les composés dans le mélange ayant provoqué les variations de déplacements chimiques<sup>72</sup>. Dû à la grande symétrie présente dans la DHFR R67 et au fait que le site potentiel de liaison des ligands se trouve à l'intersection des axes de symétries, cette technique n'est pas adéquate. Suite à un événement de liaison, la symétrie est perdue, ainsi, pour un même résidu (un pour chaque monomère) il y aurait quatre variations de déplacement chimiques différents, ce qui rend l'interprétation des résultats très fastidieuses.

La cristallographie par diffraction des rayons X peut aussi servir au criblage de composés<sup>73</sup>. Cette technique consiste à obtenir des cristaux de protéine et d'acquérir leur densité électronique par la diffraction des rayons X afin d'en déterminer la structure tridimensionnelle. Normalement, un mélange de plusieurs composés (jusqu'à une centaine<sup>76</sup>) est placé en présence des cristaux de protéine (*soaking*). Les cristaux sont analysés. Si la résolution le permet, les composés liant la protéine sont directement identifiés lors de cette étape, en contraste avec la RMN<sup>77</sup>. Toutefois il peut être difficile d'obtenir les cristaux de protéines. De plus, l'addition de ligands potentiels en solution peut casser ou même dissoudre les cristaux sans compter que les ligands n'ont parfois pas accès à leur site de liaison puisque la cible est sous forme cristalline. Donc plusieurs conditions doivent être réunies afin d'en faire un criblage à haut débit. Comme pour la RMN, de hautes concentrations de protéine et de composés sont requises ce qui peut



amener des problèmes de disponibilité, coût et solubilité. Aussi, il n'est pas possible d'obtenir des données thermodynamiques avec cette technique : on obtient des informations qualitatives et non quantitatives sur l'affinité. Jusqu'à maintenant, l'obtention de cristaux de la DHFR R67 est fastidieuse et les cristaux semblent fragiles lorsque mis en présence de composés chimiques (Damien Colin, Brahm Yachnin, Albert Berghuis, résultats non publiés). Pour l'instant, cette technique ne pourrait pas être appliquée à un criblage à haut débit pour la DHFR R67.

En contraste avec les techniques de RMN et de cristallographie, la titration calorimétrique (ITC) ne fournit aucune information structurale, mais plutôt des paramètres thermodynamiques : la constante d'équilibre ( $K_{eq}$ ), la stœchiométrie (N), l'enthalpie ( $\Delta H$ ), l'entropie ( $\Delta S$ ) et l'énergie libre de Gibbs ( $\Delta G$ ). Cette technique consiste à placer la cible en présence d'une quantité croissante d'un composé d'intérêt et de mesurer la quantité d'énergie dégagée (réaction exothermique) ou consommée (réaction endothermique) lors de la liaison. Si le composé interagit avec la cible, il y aura formation/bris de liaisons et un dégagement ou une consommation d'énergie sera mesurée. Toutefois cette technique est applicable uniquement aux réactions réversibles. On ne pourra pas analyser l'effet d'un inhibiteur covalent et on ne peut analyser qu'un seul composé à la fois, ce qui rend la méthode difficilement applicable à un criblage à haut débit. C'est pourquoi on ne s'est pas arrêté sur cette technique pour cribler la DHFR R67.

La résonance des plasmons de surface est de plus en plus utilisée dans le criblage de composés<sup>75</sup>. Cette méthode consiste à immobiliser la cible sur une couche mince de métal comme de l'or ou de l'argent. Étant donné qu'un métal possède beaucoup d'électrons, son nuage électronique (plasmon) est facilement déformable. Il existe une oscillation de ce nuage qui entoure les atomes de métal et cette oscillation possède une fréquence précise. Lorsqu'un photon possède une fréquence semblable à la fréquence d'oscillation du nuage électronique (longueur d'onde d'excitation), il y aura un effet de résonance. Le photon sera alors absorbé puis le plasmon excité. Lorsqu'une molécule se lie à la surface du métal, changeant sa masse, cela induit un changement des propriétés

du plasmon. La fréquence de résonance sera modifiée et donc la fréquence d'excitation le sera aussi. Cette modification de la longueur d'onde d'excitation peut être observée à l'aide d'un spectrophotomètre. Cette technique est facilement applicable au criblage à haut débit et ne nécessite que peu de composés/protéines<sup>75</sup>. Par contre, il faut immobiliser la cible sur la surface de métal ce qui peut modifier sa structure et ainsi sa liaison avec les composés par rapport à la protéine libre. La nature tétramérique de la DHFR R67 rend cette technique difficilement applicable. Le fait de fixer les monomères de la DHFR R67 sur une surface solide pourrait empêcher la formation de la structure quaternaire.

Les techniques présentées ci-dessus ne servent qu'à déterminer la liaison d'un composé avec la cible. Elles ne tiennent pas compte de l'effet de la liaison d'un composé sur la fonction de la cible, c'est-à-dire l'effet biologique du composé. Par conséquent, il est nécessaire de développer un essai d'activité afin d'identifier l'effet des ligands. Dans le cas où le test d'activité est suffisamment rapide et automatisable, comme lorsque l'on mesure la disparition d'un substrat ou de l'apparition d'un produit par spectrophotométrie, il peut être utilisé lors du criblage. Ceci permet de découvrir directement des molécules ayant l'effet recherché sur la cible isolée ou encore sur des cellules en culture. Dans le cas où le test d'activité est trop complexe, on pourra se pencher sur l'une des méthodes mentionnées plus haut afin d'identifier des molécules candidates. Dans le cas de la DHFR R67, l'essai d'activité est simple et applicable à un criblage à haut débit. Le développement de cet essai est expliqué en détail dans le chapitre 2.

Lorsqu'on développe un essai, il est important qu'il soit sensible et reproductible. Une approche utilisée afin de déterminer quantitativement la robustesse d'un essai est le "Zscore". Le Zscore prend en compte la reproductibilité de l'essai (écart-type du contrôle négatif et positif) et sa sensibilité (différence de signal entre le contrôle négatif et positif). Selon la formule (1), plus l'écart-type est petit et plus la différence entre le contrôle négatif et positif est grande, plus le Zscore est grand. Un Zscore plus grand que

0.5 est considéré acceptable et démontre que l'essai est fiable. Un Zscore en dessous de 0.5 démontre au contraire que l'essai n'est pas adéquat<sup>78</sup>.

$$\text{Zscore} = 1 - \left( \frac{(3 \times \sigma^-) + (3 \times \sigma^+)}{(\mu^+ - \mu^-)} \right) \quad (1)$$

$\sigma^-$  = Écart-type du contrôle négatif

$\sigma^+$  = Écart-type du contrôle positif

$\mu^-$  = Moyenne du signal du contrôle négatif

$\mu^+$  = Moyenne du signal du contrôle positif

Le développement de l'essai guidera le choix de l'approche à prendre dans la recherche de composés. S'il peut être automatisé, il sera possible de faire un criblage à haut débit (highthroughput screening ; HTS), méthode brute qui consiste au criblage *in vitro* de la cible isolée ou en culture cellulaire avec des banques de composés pouvant contenir jusqu'à quelques millions de composés. Le HTS repose essentiellement sur le fait que plus le nombre de composés criblés est important, plus les chances d'identifier des molécules actives augmentent. Cette technique a souvent été utilisée avec succès dans le développement de médicaments<sup>79</sup>. Par contre, beaucoup de programmes de HTS ont été arrêtés une fois arrivé à l'étape d'optimisation<sup>80</sup>. En effet, les composés actifs identifiés de cette façon sont généralement loin d'être optimaux. Les molécules forment des liaisons sub-optimales et possèdent souvent des groupements indésirables qui provoquent une déstabilisation lors de la liaison du ligand sur la cible. En outre, ces composés ont souvent des propriétés physicochimiques donnant lieu à une mauvaise perméabilité ou une instabilité métabolique. Bien que les librairies utilisées en HTS semblent énormes, elles ne représentent qu'une faible fraction de toutes les  $\approx 10^{60}$  molécules « *drug-like* » possédant une masse molaire d'environ 500 g/mol<sup>81</sup>. Ainsi, une campagne de criblage HTS ne couvre qu'une infime partie de l'espace chimique connu.

De plus, la diversité chimique de ces banques est relativement restreinte, donc l'espace chimique criblé n'est pas très large. Quand on prend conscience de toutes les exigences qu'une molécule doit remplir pour devenir un bon médicament et du nombre de molécules possibles, on comprend qu'une méthode brute n'est pas optimale<sup>82</sup>.

Un moyen de couvrir un plus grand espace chimique est le design à base de fragments (voir la section 1.3.1 pour plus de détails). Cette approche consiste à tester des composés ayant un faible poids moléculaire (environ 150 à 250 g/mol). Les fragments actifs seront ensuite liés pour former une molécule plus volumineuse et puissante. On estime qu'il existe environ 44 millions de composés de 160 g/mol<sup>81</sup>, ce qui est nettement moins que le nombre estimé de  $10^{60}$  composés ayant une masse molaire d'environ 500 g/mol. Il est estimé que le criblage de 1000 fragments couvre le même espace chimique que celui d'un million de composés retrouvés dans une banque typique de HTS<sup>83</sup>. La logique derrière cette estimation est la suivante : si on imagine une molécule de 450 g/mol qui correspond à 3 fragments de 150g/mol, et si chaque fragment possède 5 possibilités de se lier entre eux, alors on doit synthétiser et tester  $5^3$  molécules de 450 g/mol pour couvrir l'espace chimique de 3 fragments de 150 g/mol<sup>82</sup>. Ainsi, une librairie de 1000 à 10 000 fragments couvre une fraction non-négligeable de l'espace chimique connu, au contraire du HTS classique<sup>82</sup>. Le taux de succès du design à base de fragments semble de 10 à 1000 fois meilleur que celui du HTS<sup>79</sup>. Pour découvrir des inhibiteurs de la DHFR R67, l'approche de design par fragments a été retenue dû au fait qu'elle ne nécessite pas de grandes librairies de composés complexes et qu'elle est plus efficace pour cribler l'espace chimique.

Le criblage par arrimage moléculaire (*molecular docking*) est une technique *in silico* qui est de plus en plus utilisée depuis environ 10 ans<sup>84</sup>. Cette technique s'apparente au HTS vu que les deux méthodes sont des méthodes brutes, c'est-à-dire que l'on teste le plus grand nombre de composés possibles afin d'identifier des composés actifs. Tel qu'expliqué ci-dessus, en HTS, les composés sont testés *in vitro*, tandis qu'en arrimage moléculaire, les composés sont testés de manière *in silico*, c'est-à-dire de façon virtuelle. Des méthodes informatiques permettent de modéliser la liaison des composés sur la

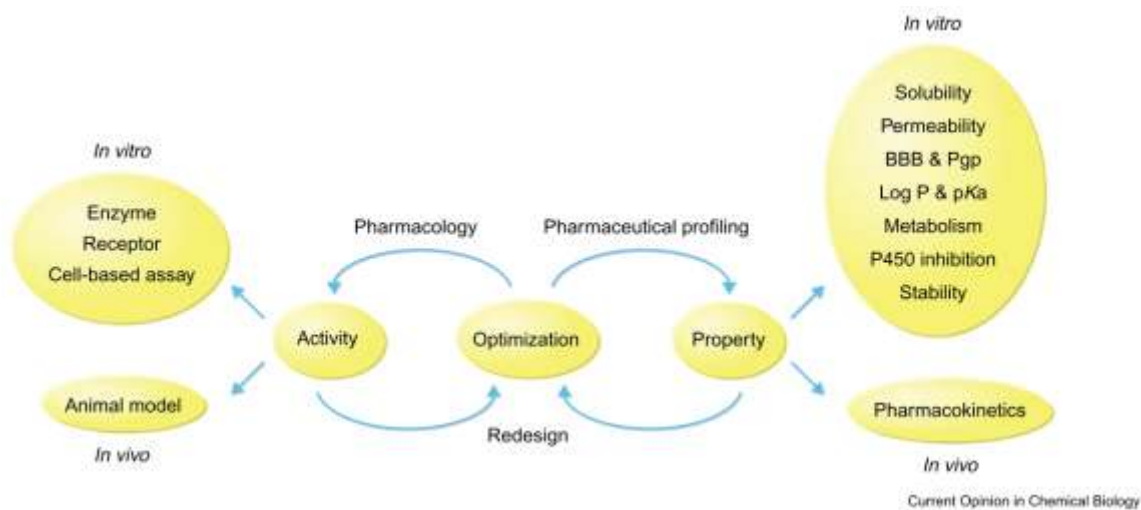
cible d'intérêt. Ces modélisations donnent une approximation de l'énergie de liaison sous la forme d'un score (voir section 1.3.2 pour plus de détails) qui permet de choisir les meilleurs candidats pour les tests *in vitro*. Ainsi, au lieu de tester des millions de composés *in vitro*, on en teste un nombre restreint, présélectionné *in silico*.

Un des avantages du criblage virtuel est la possibilité de tester à faible coûts de nombreuses librairies différentes, permettant de cribler un plus grand espace chimique, au lieu d'acheter des librairies physiques coûteuses. Par exemple, la base de données ZINC<sup>85</sup> ([www.zinc.docking.org](http://www.zinc.docking.org), visité le 28 mars 2012) permet l'accès gratuit aux banques virtuelles de plus de 180 fournisseurs commerciaux totalisant plus de 14 millions de composés<sup>86</sup>. Ainsi, les laboratoires ne sont plus limités à leurs banques de composés internes. Il faut noter que le criblage virtuel est indépendant des conditions *in vitro*. Cela peut être intéressant dans le cas où il n'est pas possible de mettre au point un essai automatisable pour la cible d'intérêt. Aussi, le criblage virtuel donne des informations au niveau de la liaison du composé sur la cible. Il est possible de visualiser les liaisons et les atomes impliqués dans la liaison du composé sur la cible. Évidemment, la plus grande faiblesse d'un tel criblage se situe au niveau de la justesse des prédictions faites par les logiciels. Il est difficile d'évaluer le taux de succès du criblage virtuel dû aux nombreuses méthodes disponibles<sup>87</sup> et du fait que seuls les résultats positifs sont rapportés<sup>84</sup>. En l'an 2000, on répertoriait environ 5 publications sur le criblage virtuel tandis qu'en 2009, plus de 100 publications ont été rapportées. Ce qui démontre un intérêt croissant et le perfectionnement de cette approche. Un autre désavantage de cette approche est qu'il faut que la structure tridimensionnelle de la cible soit disponible, ce qui peut être problématique dû aux difficultés que peuvent offrir certaines cibles pour obtenir leurs structures.

Cette technique a aussi été retenue pour découvrir des inhibiteurs de la DHFR R67 puisque plusieurs structures tridimensionnelles de grande qualité sont disponibles. De plus, cribler des composés plus complexes est complémentaire au design par fragments qui lui, se concentre surtout sur de petits composés simples. Ainsi en utilisant en parallèle le design par fragment ainsi que le criblage virtuel, on maximise les chances de découvrir des inhibiteurs de la DHFR R67.

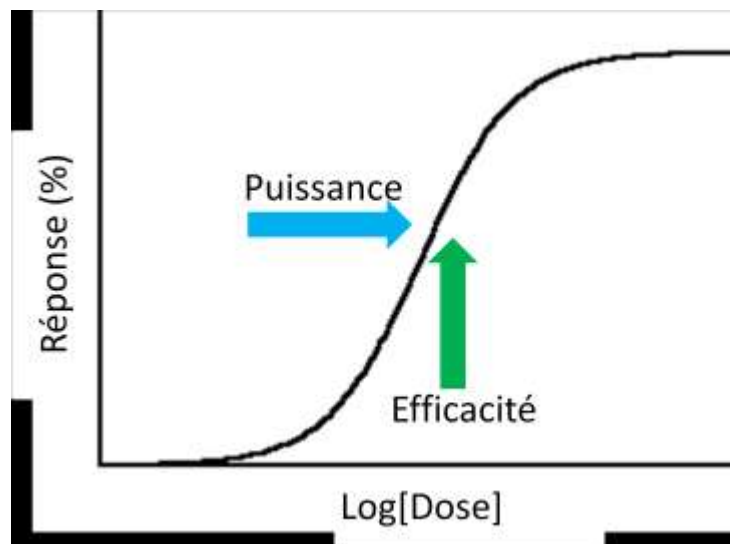
### Section 1.3.3 Détermination et caractérisation de la tête de série

Le criblage d'une cible donnera généralement lieu à l'identification de plusieurs composés actifs. Il est toutefois nécessaire de corroborer leur effet pour éliminer les faux positifs (figure 1.10e). La première étape consiste à confirmer l'intégrité du composé et sa pureté. Ceci peut être fait par chromatographie liquide couplée à la spectrométrie de masse ou à la RMN<sup>88</sup>. Suivra l'étape du profilage, lors duquel on évalue les propriétés du composé au niveau de son activité et de ses propriétés médicamenteuses (figure 1.12<sup>89</sup>).



**Figure 1.12: Évaluation et optimisation parallèle de l'activité et des propriétés médicamenteuses des composés actifs.** BBB = blood brain barrier (barrière sang-cerveau), Pgp = P-glycoprotein, qui sont des transporteurs d'efflux, c'est-à-dire, ils transportent leurs substrats de l'intérieur de la cellule vers l'extérieur. La figure est tirée de la référence 88.

Une courbe dose-réponse sera construite afin de déterminer la puissance et l'efficacité du composé (figure 1.13). La concentration effective ( $EC_{50}$ ), soit la concentration de composé nécessaire pour atteindre 50 % de la réponse d'inhibition, peut alors être mesurée.



**Figure 1.13: Courbe dose-réponse.** La puissance est évaluée selon la concentration de ligand nécessaire pour induire une réponse. L'efficacité correspond à l'amplitude de la réponse.

La sélectivité des composés actifs sera ensuite déterminée en comparant l'effet d'un composé sur la cible et sur des molécules apparentées. Il est possible que certains composés aient seulement une activité apparente et que cette activité provienne d'un autre phénomène physique. En effet, des composés peuvent former des agrégats qui absorbent la cible et en réduisent la concentration dans l'essai, donnant lieu à une inhibition apparente<sup>90</sup>. Il n'est pas possible d'établir une relation structure fonction (SAR) avec des inhibiteurs ubiquistes, car ils possèdent un mode d'action général. Il est donc important de tester les composés d'intérêt sur d'autres protéines apparentées et non-apparentées afin de déterminer si les composés provoquent un effet spécifique à la cible.

Un autre but du profilage est d'établir une relation structure-activité (SAR). À cet effet, on doit être en mesure de moduler l'activité des composés en modifiant leur structure pour en déterminer les groupements importants. Les chimistes médicaux créeront un grand nombre d'analogues afin d'être en mesure d'observer des corrélations entre la structure des composés et leur effet sur la cible. Les techniques de cristallographie, arrimage moléculaire ou de RMN aideront à déterminer les interactions

que les composés et leurs analogues établissent avec la cible pour guider les efforts de synthèse et les éventuelles étapes d'optimisation.

Environ 40 % des composés dont le développement s'est arrêté avant le stade de médicament au cours des années 2000 ont échoué à cause de leurs pauvres propriétés médicamenteuses<sup>91</sup>. Ainsi, les compagnies pharmaceutiques déterminent de plus en plus tôt les propriétés médicamenteuses de leurs composés afin d'augmenter le taux de succès<sup>92</sup>. Les propriétés étudiées comprennent la solubilité, la perméabilité, la lipophilicité, la stabilité et l'inhibition des cytochromes P450. Ces propriétés influencent le profil ADME (absorption, distribution, métabolisation et excrétion). Plus précisément, la solubilité, la lipophilicité et la perméabilité peuvent influencer l'absorption et la distribution par le biais des pompes à efflux (protéines ABC) ou à influx (protéine SLC). La métabolisation peut être influencée par la stabilité du composé et l'inhibition des cytochromes P450 par l'action directe de ce dernier ou de ses métabolites. Ainsi, en tenant compte des propriétés médicamenteuses et de l'activité de la molécule, il est possible de choisir le composé dont on tentera d'améliorer toutes les propriétés discutées plus haut : c'est la tête de série.

Au cours de nos travaux, le profilage des inhibiteurs de la DHFR R67 s'est fait en déterminant la puissance des inhibiteurs de la DHFR R67 (section 1.5.1), leur sélectivité par rapport à la DHFR humaine (section 1.5.2), l'efficacité de liaison (1.5.3), le mode d'inhibition (section 1.5.4) et les interactions impliquées entre ces molécules et la DHFR R67. Ainsi, on a été en mesure de découvrir des inhibiteurs sélectifs de la DHFR R67, d'étudier leur mode de liaison et d'identifier des groupements fonctionnels importants sur les inhibiteurs (relation structure activité).



### Section 1.3.4 - Optimisation de la tête de série et tests précliniques

L'étape d'optimisation et les tests précliniques sont difficilement dissociables car il s'agit d'un processus en boucle. Lorsqu'un composé échoue les tests précliniques, on retourne à l'étape d'optimisation pour le modifier afin d'améliorer ses propriétés.

La puissance du composé est un des aspects les plus importants car afin de passer aux tests *in vivo*, ce critère doit se trouver, au moins, dans les bas nanomolaires<sup>93</sup> et posséder une sélectivité évidente par rapport à des cibles analogues. Pour y arriver, il faut améliorer la SAR avec des expériences qui fourniront des informations structurales comme la cristallographie, la RMN ou la modélisation pour guider la modification efficace de la molécule. Une fois qu'un composé avec une puissance/sélectivité suffisantes a été identifié, on passe aux tests *in vivo*. Les paramètres de pharmacocinétique (ADME), l'effet thérapeutique, le mécanisme d'action et la toxicité du composé seront alors évalués.

Au niveau de la pharmacocinétique, il faut que le composé soit en mesure de se rendre adéquatement à son site d'action (biodisponibilité) et il doit avoir une demi-vie appropriée à l'utilisation prévue. Le composé doit montrer une efficacité évidente dans les modèles animaux.

En ce qui concerne la toxicité, la tête de série ne doit pas inhiber significativement les cytochromes P450. Il doit démontrer une absence évidente de toxicité sur des cibles clés, par exemple les hERG. Il s'agit de canaux ioniques cardiaques qui sont souvent la cible de médicaments : leur inhibition entraîne des problèmes cardiaques<sup>94</sup>. Aucun autre signe apparent de toxicité évidente ne doit être visible. Le composé ou ses métabolites ne doivent pas, par exemple, avoir de propriétés carcinogènes à court ou moyen terme (jours et mois). On continuera d'améliorer les propriétés de la tête de série afin de passer aux tests cliniques<sup>93</sup>.

### Section 1.3.5 - Essai clinique

Pour recevoir le titre de candidats médicaments et avant de pouvoir être administrée aux humains, la tête de série doit remplir plusieurs critères. Au niveau chimique, les conditions d'entreposage du composé doivent conférer au composé une bonne stabilité afin de garantir que seulement la molécule d'intérêt sera administrée aux humains et non un mélange de la molécule d'intérêt et de produits de dégradation<sup>93</sup>.

Le mécanisme d'action et le profil ADME doivent être pleinement déterminés, les voies d'absorption et d'élimination principales et secondaires doivent être clairement connues. Les voies métaboliques importantes qui pourraient influencer l'effet d'action et la dégradation du composé doivent être déterminées ainsi que ses métabolites et les interactions médicaments-médicaments doivent être établies<sup>93</sup>.

Au niveau de la toxicité, les études systématiques de tissus d'au moins deux espèces animales ne doivent pas montrer de signes apparents de toxicité. Aucun signe de carcinogénèse ou de cardiotoxicité ne doit être observable<sup>93</sup>. La dose sans effet thérapeutique, la dose thérapeutique, la dose sans effet toxique et finalement la dose létale doivent être mesurées sur les modèles animaux afin d'estimer les doses chez l'humain. Il doit y avoir une grande différence entre la dose létale et la dose thérapeutique.

Lors de la phase I, le candidat-médicament est administré à un nombre restreint de volontaires (moins d'une centaine) en bonne santé afin d'évaluer d'éventuels effets secondaires. Durant cette phase les premières données pharmacocinétiques chez l'humain sont obtenues. Le candidat médicament peut aussi être administré chez des patients pour lesquels ce traitement représente la seule chance de survie. Lors de cas exceptionnels, par exemple lors de l'étude de médicaments chimiothérapeutiques, la phase I est éliminée et ces composés passent directement en phase II : il est déjà bien connu que ce type de médicament provoque d'importants effets secondaires. Ainsi, il est inutile de tester ces composés chez des humains en santé.

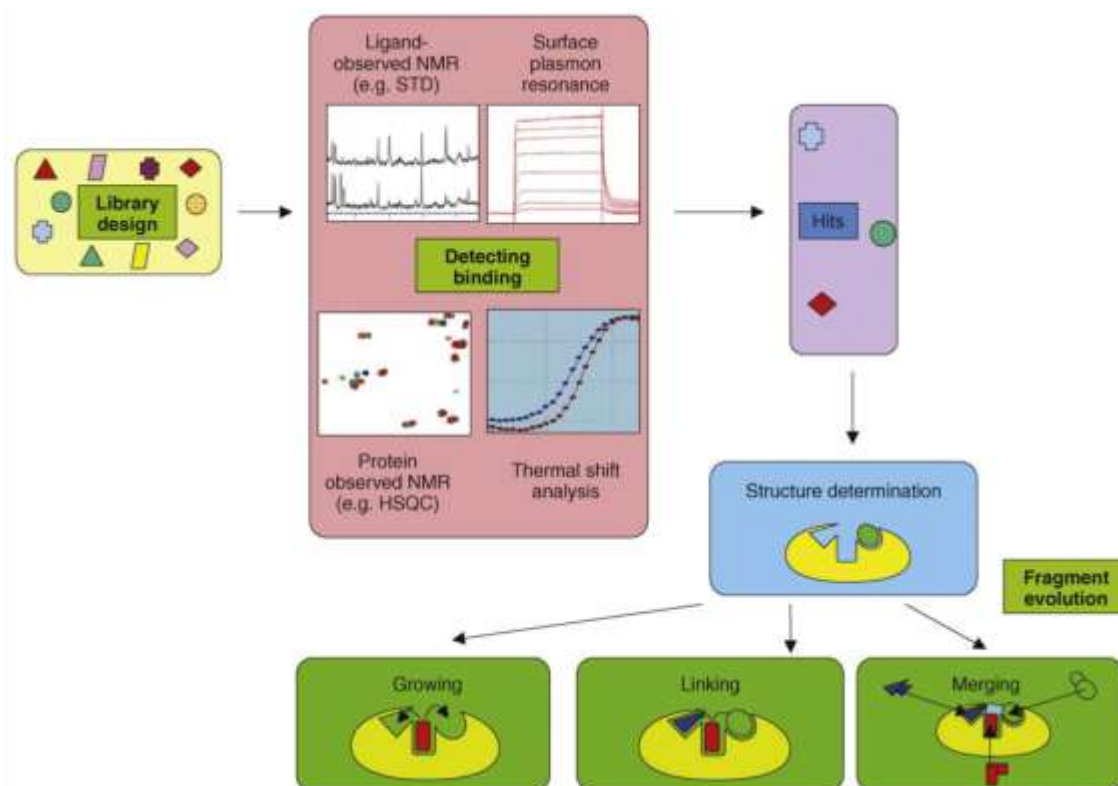
En phase clinique II on teste l'efficacité thérapeutique du candidat-médicament. On donne le composé à un groupe de volontaires plus important qu'en phase I, atteints de la pathologie d'intérêt. Au cours de cette phase, on détermine la dose thérapeutique chez l'humain et on continue d'étudier d'éventuels effets toxiques. La phase III est celle où l'efficacité du candidat-médicament est comparée à un médicament analogue (si applicable) et à un placebo. Lors de cette phase, les groupes de volontaires sont d'une taille importante ce qui permet de déceler des effets secondaires rares ou peu importants. Finalement lors de la phase IV, qui est celle de commercialisation, il y aura un suivi à long terme du médicament et des études statistiques à grande échelle sur de longues périodes de temps pourront être entreprises afin de déceler des effets secondaires très rares et tardifs.

## **Section 1.4 - Identification de composés actifs : inhibiteurs**

Pour découvrir des inhibiteurs de la DHFR R67, deux approches ont été retenues soient celle du design à base de fragments et celle de l'arrimage moléculaire. Ces deux techniques seront expliquées en détails dans les deux prochaines sous-sections.

### **Section - 1.4.1 - Design à base de fragments**

Les avantages et désavantages du design à base de fragments ont été abordés dans la section 1.2.2. Comme mentionné dans cette section, le design à base de fragments consiste à découvrir des petites molécules simples qui inhibent faiblement la cible (ordre du millimolaire) pour ensuite construire une molécule plus grosse et plus puissante (figure 1.14<sup>95</sup>).



**Figure 1.14 : Processus typique de design à base de fragments.** Une librairie de petites molécules est criblée sur la cible. Des composés actifs sont découverts. Plusieurs petites molécules sont liées ensemble afin de former une plus grosse molécule plus puissante. La figure est tirée de la référence 94.

Une librairie de fragments est criblée sur une cible. Des fragments qui interagissent avec la cible sont découverts. Bien que les fragments forment moins d'interactions avec la cible et sont moins puissants que leurs homologues découverts en HTS, leurs interactions avec la cible sont de plus grande qualité<sup>83</sup>. Cette notion peut se comprendre avec les équations 2 et 3.

$$\Delta G = \Delta H - T\Delta S \quad (2)$$

$\Delta G$  = Énergie libre de Gibbs

$\Delta H$  = Enthalpie

T = Température en Kelvin

$\Delta S$  = Entropie

$$\Delta G = RT \ln K_i \quad (3)$$

R = constante des gaz parfait

$K_i$  = constante d'inhibition

Toute molécule qui se lie à une cible doit payer une pénalité en énergie sous la forme d'entropie. En effet, quand une molécule se lie sur une cible, elle perd ses entropies de translation et de rotation qu'elle possède en solution<sup>96</sup>. Il est estimé que cette pénalité est d'environ  $\Delta S = +15\text{-}20$  kJ/mol pour toutes molécules, peu importe leur grosseur<sup>97</sup>. Ainsi, un fragment qui possède un  $K_i$  de 100  $\mu\text{M}$ , ce qui représente un  $\Delta G$  de -22.8 kJ/mol, possède en réalité un  $\Delta H$  de -40 kJ/mol (interaction favorable). Un inhibiteur ayant un  $K_i$  de 3 nM possède un  $\Delta G$  d'environ -60 kJ/mol, donc il doit avoir un  $\Delta H$  de -80 kJ/mol (interaction favorable). Ainsi, le fragment avec son  $\Delta H = -40$  kJ/mol qui ferait partie de ce puissant inhibiteur, fournirait à lui seul la moitié de l'énergie de liaison, à condition que ses interactions avec la protéine restent les mêmes une fois incorporé dans la plus grosse molécule<sup>83</sup>.

Justement, le plus grand défi dans le design à base de fragments est de faire évoluer les fragments en gardant leurs interactions avec la cible. Pour ce faire, 3 approches ont été développées : il s'agit de la croissance du fragment, la liaison des fragments et la fusion des fragments (figure 1.14).

La croissance des fragments consiste à améliorer la puissance du fragment en ajoutant des groupes fonctionnels à un ou plusieurs endroits sur la molécule. Par conséquent, dans cette approche, le fragment joue plutôt le rôle d'une ancre<sup>98</sup>. Des informations structurales (RMN ou cristallographie) sont nécessaires pour savoir comment ajouter les groupements sur la molécule afin de créer de nouvelles interactions avec la protéine. Dû à son rôle d'ancre, le fragment ne doit pas changer de mode de liaison pendant le processus de croissance vers un inhibiteur plus puissant<sup>98</sup>. Cette approche a été récemment utilisée avec succès dans le développement d'inhibiteurs pour BACE-1 ( $\beta$ -sécrétase)<sup>99</sup>. Dans l'industrie pharmaceutique, cette enzyme est considérée

comme une cible très difficile, mais son rôle biologique dans le développement de l'Alzheimer la rend très intéressante<sup>98</sup>. L'approche de croissance de fragments a été aussi utilisée avec succès pour le développement d'inhibiteurs contre l'urokinase<sup>100</sup> et la phosphodiesterase 4<sup>101</sup>.

L'approche de la liaison de fragments consiste à identifier plusieurs fragments qui se lient sur la cible et les lier ensemble directement ou avec des espaceurs<sup>98</sup>. Cette manière de procéder est potentiellement plus rapide que la croissance des fragments dans le but d'obtenir un composé puissant. Par contre, le grand défi est d'être en mesure de lier les différents fragments ensemble sans changer leurs mode de liaison<sup>98</sup>. Des informations structurales sont un atout, mais ne sont pas essentielles. Récemment, cette approche a été utilisée dans le développement d'inhibiteurs pour Bcl-X<sub>L</sub> qui est une cible dans le combat du cancer<sup>102</sup>.

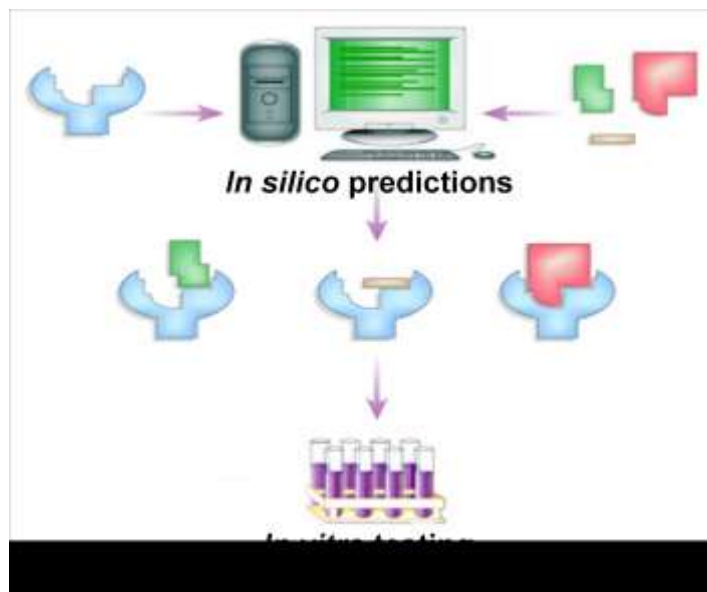
La dernière approche est celle de fusion des fragments qui est une technique *in situ*. Cette méthode est basée sur la chimie « *click* » pour lier les fragments entre eux. On ajoute des groupements réactionnels aux fragments qui pourront réagir entre eux lorsque liés sur la cible afin de former des liens covalents et créer une plus grande molécule<sup>103</sup>. Ensuite, cette plus volumineuse molécule pourra être analysée et identifiée. Les conditions de réaction doivent être douces et se faire dans un environnement aqueux à température ambiante ce qui est possible dans le cas de la chimie « *click* »<sup>98</sup>. L'avantage de cette technique est qu'elle ne nécessite pas d'avoir d'information structurale. Cette approche a été utilisée dans le développement d'inhibiteurs contre l'anhydrase carbonique<sup>104</sup>.

Ainsi, le design à base de fragments permet de cribler plus efficacement l'espace chimique connu, ce qui a permis d'identifier des inhibiteurs pour des cibles difficiles. De plus, il permet d'identifier des molécules qui forment des interactions de grande qualité avec la cible, ce qui rend le processus d'optimisation de la puissance plus rapide. Finalement, l'étape de criblage est plus simple, car les librairies de fragments sont généralement plus petites que les librairies de HTS. La conception d'une librairie de

fragments est plus facile dû au fait que ces composés sont souvent simples et disponibles commercialement, contrairement aux composés retrouvés dans les librairies de HTS qui nécessitent des synthèses complexes. Pour ces raisons, le choix du design par fragment pour la recherche d'inhibiteur de la DHFR R67 a été entrepris (chapitre 3).

### Section - 1.4.2 - Arrimage moléculaire

Les avantages et désavantages du criblage par arrimage moléculaire ont été discutés à la section 1.3.2. Comme il a été mentionné à cette section, l'arrimage moléculaire est une technique *in silico* où des modèles de liaison de ligands sur une protéine sont créés afin de prédire leurs interaction (énergie de liaison et pose) avec la protéine d'intérêt (figure 1.15<sup>105</sup>).

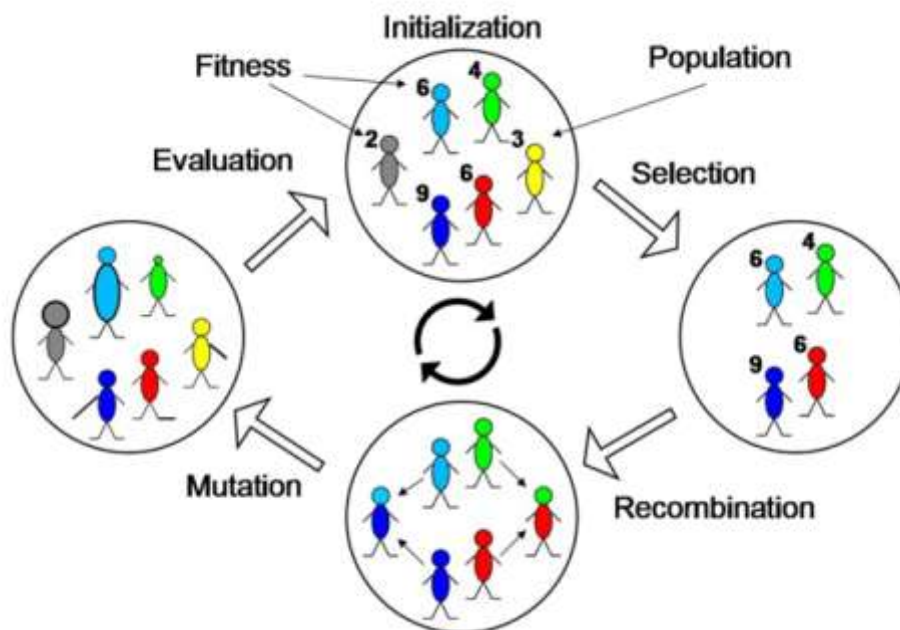


**Figure 1.15 : Processus typique de criblage par arrimage moléculaire.** Une librairie virtuelle est criblée sur la cible. Le logiciel prédit la liaison potentielle de certains composés. Ensuite, ces composés sont testés *in vitro*. La figure est tirée de la référence 103.

Les logiciels d'arrimage moléculaire génèrent plusieurs milliers de poses différentes d'un composé à l'intérieur du site de liaison de la cible et évaluent chaque pose selon un score qui est une estimation de l'énergie de liaison. Cette procédure peut être faite sur des millions de composés et ensuite, selon le score, il est possible de déterminer les meilleurs candidats pour les tests *in vitro*.

Ces logiciels fonctionnent selon deux étapes principales, celles d'arrimage (génération d'une pose) et de détermination du score (prédiction de l'énergie de liaison). Pour un composé possédant 4 rotations libres, il existe  $1.7 \times 10^{10}$  ( $360^4$ ) poses possibles en émettant l'hypothèse que chaque rotation tourne d'un degré. Pour un composé possédant seulement une rotation libre de plus, le nombre de poses passe à  $6 \times 10^{12}$  ( $360^5$ ). Ensuite, on peut imaginer que le composé entier peut subir des rotations sur l'axe X, Y et Z, ce qui multiplie le nombre de poses possibles par un facteur de  $4.7 \times 10^7$  ( $360^3$ ) en supposant encore une fois une rotation d'un degré au minimum. Finalement, ce composé peut aussi subir des mouvements de translation dans l'espace. On peut alors comprendre que même les ordinateurs les plus puissants ne pourront jamais générer et évaluer autant de poses pour trouver la meilleure pose parmi toutes les poses possibles (excluant la translation du composé), et ce, pour un seul composé en un temps raisonnable. Ainsi, une approche par force brute n'est pas adéquate. Donc pour rechercher efficacement la meilleure pose parmi autant de poses possibles, des algorithmes d'arrimage ont été développés. Un algorithme d'arrimage peut utiliser l'échantillonnage systématique pour accélérer sa recherche<sup>106</sup>. Au lieu d'évaluer systématiquement chaque pose créée, il évalue seulement une pose à chaque tranche d'un nombre déterminée; ainsi, les poses similaires ne seront pas évaluées, ce qui diminue la charge de calcul. Les algorithmes génétiques correspondent à un autre type d'algorithme développé pour la recherche de poses. Ces algorithmes sont basés sur la théorie de l'évolution de Darwin. Une population (poses) évolue et transmet les gènes favorables (propriétés de pose favorables) aux descendants (nouvelles poses) et les gènes défavorables (propriétés de pose défavorable) sont éliminés<sup>106</sup> (figure 1.16<sup>107</sup>).





**Figure 1.16: Schématisation du fonctionnement d'un algorithme génétique.** Les explications se trouvent dans le texte. La figure est tirée de la référence 104.

À l'étape d'initialisation, l'algorithme génère aléatoirement une population de départ. Chaque pose est ensuite évaluée selon l'algorithme de score (voir ci-dessous). À l'étape de sélection, les poses ayant le meilleur score sont gardées et les scores faibles sont éliminées. À l'étape de recombinaison, afin de remettre à niveau la population, des poses « enfants » sont créées à partir de segments des poses « parents » ce qui donne des hybrides entre les meilleures poses. À l'étape de mutation, des rotations aléatoires sont incluses dans certaines poses de la population. Finalement, à la phase d'évaluation, la population est évaluée et le processus recommence. Ainsi, l'algorithme converge vers la meilleure pose théorique.

Pour évaluer les poses, les logiciels d'arrimage moléculaire utilisent des algorithmes de score. Ces algorithmes estiment l'énergie de liaison de plusieurs façons différentes. Les fonctions de score basées sur les champs de force reposent sur la physique atomique en tenant compte des interactions Van der Waals, des interactions électrostatiques, des forces de torsion, de rotation et enfin d'élongation des liens<sup>108</sup>. Ces fonctions sont normalement dérivées de données expérimentales et des connaissances

théoriques en physique quantique<sup>109</sup>. L'équation 4<sup>109</sup> montre une équation typique de fonctions de score de champ de force qui met en relation les forces mentionnées.

$$E = \sum_i \sum_j \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} \right) \quad (4)$$

$r_{ij}$  = distance entre atome i de la protéine et l'atome j du ligand

$A_{ij}$  et  $B_{ij}$  = paramètre van der Waals

$q_i$  et  $q_j$  = charge des atomes i et j

$\epsilon(r_{ij})$  = constante diélectrique pour l'effet du solvant

Ce type de fonction peut donner des estimations précises pour prédire l'affinité d'un ligand avec la cible ( $K_i$ ) mais à cause de ses paramètres complexes, ce type de fonction est informatiquement lourde.

Les fonctions empiriques correspondent à un autre type de fonctions de score. Ce type de fonction estime l'énergie de liaisons entre un ligand et une cible d'après un ensemble de termes énergétiques prédéfinis. Par exemple, on pourrait assigner à la formation d'un pont H une variation de -5 sur le score et aux interactions de Van der Waals, qui sont plus faibles, un changement seulement de -2 sur le score. Ainsi, tous les thèmes énergétiques tels que les interactions Van der Waals, la formation de ponts H, les interactions électrostatiques, la désolvatation, l'entropie, l'hydrophobicité peuvent être prédéfinis. L'équation 5<sup>109</sup> montre une équation typique d'une fonction empirique.

$$\Delta G = \sum_i W_i \cdot \Delta G_i \quad (5)$$

$\Delta G_i$  = énergie assignée aux différents termes énergétiques

$W_i$  = constante théorique basée sur l'entraînement de la fonction

L'avantage de ce type de fonction par rapport aux autres basées sur les champs de force est qu'il est beaucoup plus rapide dû à ses thèmes énergétiques plus simples<sup>109</sup>.

Grâce à ces derniers, ces fonctions ne visent pas à déterminer l'affinité absolue des ligands (prédictions d'un  $K_i$ ), mais plutôt à assigner une affinité relative afin de déterminer quels composés, parmi une librairie, auraient le plus de probabilités de se lier à la cible. Donc elles sont rapides mais peu précises.

Une autre fonction de score est constituée par les fonctions basées sur les connaissances (knowledge based scoring functions). Les termes énergétiques retrouvés dans ces fonctions sont dérivés d'études statistiques sur des bases de données d'informations structurales obtenues expérimentalement. L'énergie d'interaction entre les atomes d'une cible et d'un composé est estimée sur des interactions similaires qui ont déjà été auparavant mesurées expérimentalement. Par exemple, si un logiciel utilisant cette fonction de score détermine la formation d'un pont H entre un composé et la cible d'intérêt, une recherche dans une base donnée (comme celle de Cambridge Structural Database<sup>110</sup>) sera faite pour trouver des ponts H correspondant à celui déterminé (atomes impliqués, distance, torsion, etc) et une valeur moyenne d'énergie de liaison lui sera attribuée. Ainsi, ce type de fonction présente un compromis entre la vitesse et la précision<sup>109</sup>. Par rapport aux fonctions empiriques, on gagne en précision dû au fait que les termes énergétiques sont basés sur des mesures expérimentales. On gagne en vitesse par rapport aux fonctions de champ de force dû au fait que les termes énergétiques ne sont pas aussi complexes.

Plusieurs campagnes de criblage par arrimage moléculaire ont été entreprises avec succès. Seulement dans *Journal of Medicinal Chemistry*, 141 rapports au sujet de la découverte de composés actifs par arrimage moléculaire ont été publiés<sup>84</sup>. Dans 60 de ces publications, des composés ont été rapportés avec une affinité en dessous de 1  $\mu\text{M}$ . 55 publications ont rapporté des composés avec une affinité entre 1 et 10  $\mu\text{M}$ <sup>84</sup>. Les composés découverts par arrimage moléculaire peuvent avoir le potentiel d'aller loin dans le processus de développement des médicaments. L'Aggrastat est un médicament maintenant sur le marché qui a été développé par Merck<sup>111</sup>. C'est un antagoniste des récepteurs à fibrinogène pour traiter les infarctus. Le premier composé identifié qui mena à l'Aggrastat avait un  $\text{IC}_{50}$  de 27  $\mu\text{M}$  et a été identifié par arrimage moléculaire.

Les informations structurales obtenues avec l'arrimage moléculaire ont été aussi utilisées afin d'optimiser ce composé<sup>111</sup>.

Le SC12267 est un inhibiteur de la dihydroorotate déhydrogénase pour le traitement de l'arthrite. Le projet a commencé par un criblage virtuel sur une structure cristalline de la dihydroorotate déhydrogénase et une série de composés a été identifiée<sup>112</sup>. En moins de 28 mois, le composé découvert par arrimage moléculaire est passé en test clinique<sup>113</sup>.

La pharmaceutique Novartis a annoncé la découverte du Cevoglitazar comme un agoniste du récepteur PPAR. Le Cevoglitazar a été développé à partir de composés actifs découverts par arrimage moléculaire. Le Cevoglitazar s'est rendu en phase clinique II mais a été discontinué par la suite<sup>113</sup>.

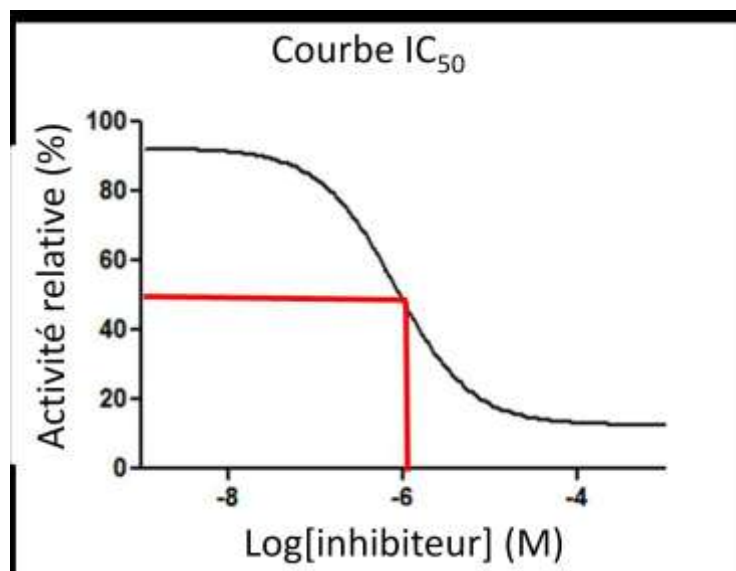
Ainsi le criblage par arrimage moléculaire permet de cribler de grandes bibliothèques de composés complexes rapidement et à faible coût. Comparé au design à base de fragments, c'est une technique qui permet l'identification de composés plus complexes pouvant démontrer une puissance plus grande dès le départ. De plus, les informations structurales provenant de l'arrimage moléculaire peuvent être utilisées dans le développement des composés. La disponibilité de plusieurs structures cristallines de DHFR R67 dans le domaine public et des banques virtuelles de composés gratuites rendent la mise au point du criblage simple. Ainsi, un criblage virtuel a été fait en parallèle avec le design à base de fragments dans le but de maximiser les chances de découvrir des inhibiteurs de la DHFR R67.

## **Section 1.5 - Profilage des inhibiteurs de la DHFR R67**

Afin de pouvoir tester les composés, un test d'activité où l'on mesure l'activité de la DHFR R67 par la consommation du substrat et du cofacteur dans le temps par spectrophotométrie a été mis au point (voir chapitre 2 pour les détails). Les composés actifs découverts ont ensuite été caractérisés selon leur puissance, sélectivité, efficacité de liaisons, mode de liaison. Des études de modélisation par arrimage moléculaire ont aussi été faites pour étudier comment les inhibiteurs interagissent avec la DHFR R67.

### Section 1.5.1 - Évaluation de la puissance par dose réponse

Pour estimer la puissance des inhibiteurs découverts lors du criblage, on évalue la concentration inhibitrice médiane ( $IC_{50}$ ) et la constante d'inhibition ( $K_i$ ). L' $IC_{50}$  correspond à la concentration d'inhibiteur requise pour inhiber la moitié de l'activité d'un essai. Il s'agit de tester l'activité pour une gamme de concentrations d'inhibiteur afin de tracer un graphique de l'activité enzymatique en fonction du log de la concentration d'inhibiteur (figure 1.17). Le point d'inflexion du graphique correspond à la valeur de l' $IC_{50}$ .



**Figure 1.17 : Graphique typique d'une courbe de  $IC_{50}$ .** En rouge, on peut voir la concentration d'inhibiteur requise pour diminuer de moitié l'activité de l'enzyme : c'est le  $IC_{50}$ .

L' $IC_{50}$  est une valeur relative qui est influencée par les conditions expérimentales comme la concentration de substrat utilisé ou la nature de ce substrat. Par exemple, pour un inhibiteur compétitif, plus la concentration de substrat utilisée dans l'essai sera grande, plus l' $IC_{50}$  sera élevée. Ainsi, pour comparer des  $IC_{50}$  entre eux, les essais doivent être faits dans les mêmes conditions expérimentales.

Afin d'obtenir une valeur absolue qui informe sur la puissance des inhibiteurs, la constante d'inhibition  $K_i$  peut être déterminée. Elle correspond à la concentration requise pour occuper la moitié des sites enzymatiques. C'est une valeur absolue qui ne dépend pas des conditions expérimentales. Le  $K_i$ , pour un inhibiteur compétitif, peut être

déterminé expérimentalement (voir section 1.4.3) ou calculé selon l'équation de Cheng and Prusoff<sup>14</sup> (équation 6).

$$K_i = \frac{IC_{50}}{1 + \frac{[substrat]}{K_M}} \quad (6)$$

Dans cette équation, pour comprendre la relation entre l'IC<sub>50</sub>, la concentration du substrat et le K<sub>M</sub>, on considère K<sub>i</sub> constant. Plus le IC<sub>50</sub> est grand, plus K<sub>M</sub> est petit et plus la concentration de substrat est grande. Ainsi, on comprend pourquoi la valeur de l'IC<sub>50</sub> dépend des valeurs expérimentales. Intuitivement, on peut imaginer qu'il faudra une plus grande concentration d'inhibiteur pour voler la place au site actif d'un bon substrat (plus bas K<sub>M</sub>) ou d'un substrat présent en grande quantité (haute concentration).

## **Section 1.5.2 - Sélectivité par rapport à la DHFR humaine**

Étant donné l'application médicale prévue pour les inhibiteurs de la DHFR R67, les inhibiteurs ont été testés sur la DHFR humaine afin de s'assurer qu'ils avaient une meilleure affinité pour la DHFR R67. Choisir des inhibiteurs ayant une sélectivité pour la DHFR R67 pourrait éviter des effets secondaires provenant de l'inhibition de la DHFR humaine et qui peuvent être importants (les inhibiteurs de DHFR humaine sont des médicaments chimiothérapeutiques, voir la section 1.3.2)

Comme mentionné à la section 1.2.3, certains composés peuvent agir comme des inhibiteurs ubiquistes. Ainsi, en testant les inhibiteurs de DHFR R67 sur la DHFR humaine, on s'assure que les composés actifs n'inhibent pas non plus l'enzyme humaine. De plus, la DHFR humaine catalyse la même réaction que la DHFR R67; ainsi, on peut s'assurer que la chute d'activité observée pour la DHFR R67 ne provient pas tout simplement d'une réaction chimique avec le DHF ou le NADPH (substrat et cofacteur, respectivement, des deux DHFR). Ainsi, le test de sélectivité avec la DHFR humaine permet de découvrir des inhibiteurs sélectifs à la DHFR R67 et sert de contrôle pour confirmer la découverte d'inhibiteurs de DHFR R67.

### Section 1.5.3 - Efficacité de liaison

Comme vu à la section 1.4.1, les fragments, malgré le fait qu'ils montrent des inhibitions faibles, forment des interactions de plus grande qualité avec la cible. Un moyen de quantifier ces interactions est de calculer l'efficacité de liaison (LE). L'efficacité de liaison peut se calculer selon l'équation 7<sup>83</sup>.

$$LE = \frac{\Delta G}{N} \quad (7)$$

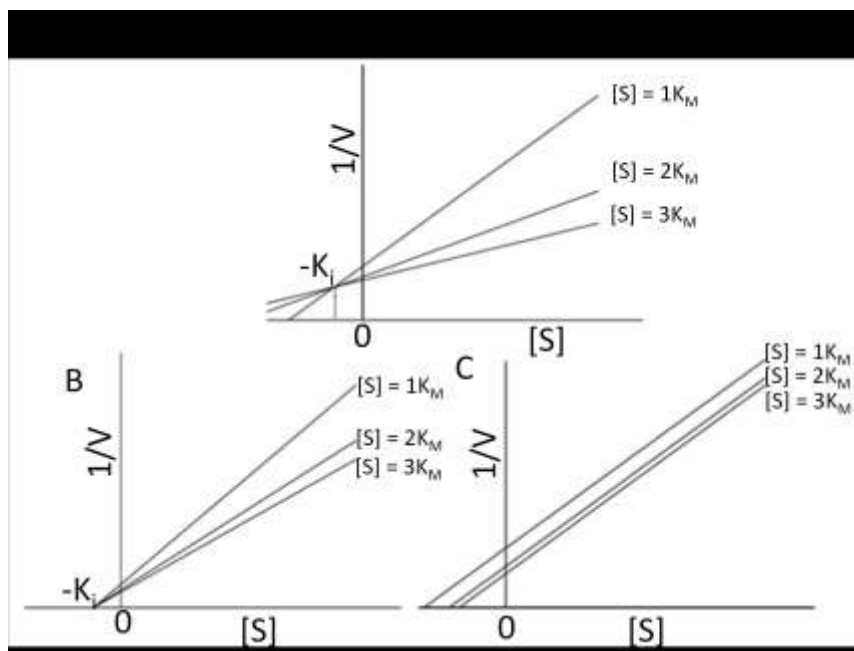
N = nombre d'atomes lourds (tous les atomes sauf les hydrogènes)

Le  $\Delta G$  peut être déterminé à partir du  $K_i$  par l'équation (3). Le  $K_i$  peut être déterminé expérimentalement ou calculé à partir de l' $IC_{50}$  par l'équation (6) si l'inhibition est compétitive. Selon l'équation 7, plus une molécule est petite (bas N) et plus elle a un bas  $\Delta G$  (donc un bas  $IC_{50}$  et  $K_i$ ), plus elle forme des interactions de qualité (bas LE). Intuitivement, on peut imaginer qu'une petite molécule ne peut pas former une grande quantité de liaisons avec la cible, donc elle doit former des liens de grande qualité pour vaincre la pénalité au niveau de l'entropie pour lui permettre de se lier à la cible. Il est estimé que des fragments ayant un LE de -0.30 kcal/mol et moins ont le potentiel de mener à des inhibiteurs de l'ordre du nanomolaire après optimisation. Donc la valeur de LE peut être utilisée pour déterminer les meilleurs fragments.

La LE peut aussi être utilisée comme un guide lors de l'optimisation des inhibiteurs. Un composé ayant une faible LE suggère que l'inhibiteur forme des liens de mauvaise qualité avec la cible ou bien qu'il possède des atomes inutiles ou nuisibles à la liaison. Donc pour obtenir une meilleure LE, on peut diminuer la taille d'un inhibiteur, ce qui peut mener à une meilleure absorption orale du même coup, ou encore, on peut modifier l'inhibiteur pour augmenter la qualité de ses interactions et sa puissance.

## Section 1.5.4 - Mode d'inhibition

Jusqu'ici, on a émis l'hypothèse que les inhibiteurs de DHFR R67 sont des inhibiteurs compétitifs car c'est le mode d'inhibition le plus commun chez les enzymes. Afin de confirmer le mode d'action des inhibiteurs de la DHFR R67, on peut tracer un graphique de Dixon<sup>115</sup>. Ce type de graphique consiste à tracer plusieurs droites de la réciproque de l'activité en fonction d'une gamme de concentrations d'inhibiteur. Chaque droite est faite à une concentration différente de substrat (voir figure 1.18). L'allure du graphique nous informera sur le mode d'inhibition. Dans un cas où les droites convergent à un point au-dessus de l'axe des X (figure 1.18A), il s'agit d'une inhibition compétitive. La valeur sur l'axe des X où les droites convergent correspond à  $-K_i$ . Dans le cas où les droites convergent sur l'axe des X (figure 1.18B), il s'agit d'une inhibition non compétitive. La valeur sur l'axe des X où les droites convergent correspond à  $-K_i$ . Dans le cas où toutes les droites sont parallèles (figure 1.18C), il s'agit d'une inhibition incompétitive. Comme les droites ne convergent pas, le  $K_i$  ne peut donc pas être déterminé dans ce cas.



**Figure 1.18 : Graphique Dixon typique pour les 3 modes d'inhibition réversibles.** A) Inhibition compétitive. B) Inhibition non compétitive. C) Inhibition incompétitive. S = concentration de substrat, V = vitesse de catalyse,  $K_M$  = constante Michaelis-Menten qui informe sur l'affinité du substrat, valeur exprimée en unité de concentration. La figure a été fournie par le professeur Jeffrey W. Keillor à l'Université d'Ottawa.



Dans le chapitre 2, nous verrons la mise au point du test d'activité et des conditions de criblages. Au chapitre 3, on verra comment ce test a été appliqué dans la découverte et la caractérisation d'inhibiteurs de la DHFR R67. Ensuite, au chapitre 4, on verra le début des études SAR. Finalement, au chapitre 5, on verra comment l'arrimage molécule a été appliqué dans la découverte d'inhibiteurs de la DHFR R67.

## Références du chapitre 1

1. Goossens, H.; Ferech, M.; Vander Stichele, R.; Elseviers, M., Outpatient antibiotic use in Europe and association with resistance: a cross-national database study. *Lancet* **2005**, *365* (9459), 579-87.
2. Bush, K.; Courvalin, P.; Dantas, G.; Davies, J.; Eisenstein, B.; Huovinen, P.; Jacoby, G. A.; Kishony, R.; Kreiswirth, B. N.; Kutter, E.; Lerner, S. A.; Levy, S.; Lewis, K.; Lomovskaya, O.; Miller, J. H.; Mobashery, S.; Piddock, L. J.; Projan, S.; Thomas, C. M.; Tomasz, A.; Tulkens, P. M.; Walsh, T. R.; Watson, J. D.; Witkowski, J.; Witte, W.; Wright, G.; Yeh, P.; Zgurskaya, H. I., Tackling antibiotic resistance. *Nat Rev Microbiol* **2011**, *9* (12), 894-6.
3. Bronzwaer, S. L.; Cars, O.; Buchholz, U.; Molstad, S.; Goettsch, W.; Veldhuijzen, I. K.; Kool, J. L.; Sprenger, M. J.; Degener, J. E., A European study on the relationship between antimicrobial use and antimicrobial resistance. *Emerg Infect Dis* **2002**, *8* (3), 278-82.
4. Ho, J. M.; Juurlink, D. N., Considerations when prescribing trimethoprim-sulfamethoxazole. *CMAJ* **2011**, *183* (16), 1851-8.
5. Forgacs, P.; Wengenack, N. L.; Hall, L.; Zimmerman, S. K.; Silverman, M. L.; Roberts, G. D., Tuberculosis and trimethoprim-sulfamethoxazole. *Antimicrob Agents Chemother* **2009**, *53* (11), 4789-93.
6. Les Burrige, J. S. W., Felipe Cabello, Jaime Pizarro, Katherine Bostick Chemical use in salmon aquaculture: A review of current practices and possible environmental effects. *Aquaculture* **2010**, *7* (23), 8 - 19.
7. Le, T. X.; Munekage, Y.; Kato, S., Antibiotic resistance in bacteria from shrimp farming in mangrove areas. *Sci Total Environ* **2005**, *349* (1-3), 95-105.
8. Kehrenberg, C.; Schwarz, S., Trimethoprim resistance in a porcine *Pasteurella aerogenes* isolate is based on a *dfrA1* gene cassette located in a partially truncated class 2 integron. *J Antimicrob Chemother* **2011**, *66* (2), 450-2.
9. Kruse, H.; Sorum, H., Transfer of multiple drug resistance plasmids between bacteria of diverse origins in natural microenvironments. *Appl Environ Microbiol* **1994**, *60* (11), 4015-21.

10. Weber, J. T.; Mintz, E. D.; Canizares, R.; Semiglia, A.; Gomez, I.; Sempertegui, R.; Davila, A.; Greene, K. D.; Puhr, N. D.; Cameron, D. N.; et al., Epidemic cholera in Ecuador: multidrug-resistance and transmission by water and seafood. *Epidemiol Infect* **1994**, *112* (1), 1-11.
11. Cabello, F. C., Heavy use of prophylactic antibiotics in aquaculture: a growing problem for human and animal health and for the environment. *Environ Microbiol* **2006**, *8* (7), 1137-44.
12. Farber, S.; Diamond, L. K., Temporary remissions in acute leukemia in children produced by folic acid antagonist, 4-aminopteroyl-glutamic acid. *N Engl J Med* **1948**, *238* (23), 787-93.
13. Rosowsky, A.; Wright, J. E.; Vaidya, C. M.; Forsch, R. A., The effect of side-chain, para-aminobenzoyl region, and B-ring modifications on dihydrofolate reductase binding, influx via the reduced folate carrier, and cytotoxicity of the potent nonpolyglutamatable antifolate N(alpha)-(4-amino-4-deoxypteroyl)-N(delta)-hemiphthaloyl-L- ornithine. *Pharmacol Ther* **2000**, *85* (3), 191-205.
14. <http://www.drugs.com/pro/methotrexate.html> (accessed january 18, 2012).
15. Hanna, N.; Shepherd, F. A.; Fossella, F. V.; Pereira, J. R.; De Marinis, F.; von Pawel, J.; Gatzemeier, U.; Tsao, T. C.; Pless, M.; Muller, T.; Lim, H. L.; Desch, C.; Szondy, K.; Gervais, R.; Shaharyar, Manegold, C.; Paul, S.; Paoletti, P.; Einhorn, L.; Bunn, P. A., Jr., Randomized phase III trial of pemetrexed versus docetaxel in patients with non-small-cell lung cancer previously treated with chemotherapy. *J Clin Oncol* **2004**, *22* (9), 1589-97.
16. Howell, E. E., Searching sequence space: two different approaches to dihydrofolate reductase catalysis. *Chembiochem* **2005**, *6* (4), 590-600.
17. Huovinen, P., Increases in rates of resistance to trimethoprim. *Clin Infect Dis* **1997**, *24 Suppl 1*, S63-6.
18. Warren, J. W.; Abrutyn, E.; Hebel, J. R.; Johnson, J. R.; Schaeffer, A. J.; Stamm, W. E., Guidelines for antimicrobial treatment of uncomplicated acute bacterial cystitis and acute pyelonephritis in women. Infectious Diseases Society of America (IDSA). *Clin Infect Dis* **1999**, *29* (4), 745-58.

19. Schneider, P.; Hawser, S.; Islam, K., Iclaprim, a novel diaminopyrimidine with potent activity on trimethoprim sensitive and resistant bacteria. *Bioorg Med Chem Lett* **2003**, *13* (23), 4217-21.
20. Hawser, S.; Lociuero, S.; Islam, K., Dihydrofolate reductase inhibitors as antibacterial agents. *Biochem Pharmacol* **2006**, *71* (7), 941-8.
21. Sader, H. S.; Fritsche, T. R.; Jones, R. N., Potency and bactericidal activity of iclaprim against recent clinical gram-positive isolates. *Antimicrob Agents Chemother* **2009**, *53* (5), 2171-5.
22. Stokstad, E. L. R., Folic acid metabolism in health and disease. In *M. F. Picciano & E. L. R. Stokstad (Eds.). New York: Wiley-Liss* **1990**, 1-21.
23. Zhao, R.; Goldman, I. D., Resistance to antifolates. *Oncogene* **2003**, *22* (47), 7431-57.
24. Moscow, J. A.; Gong, M.; He, R.; Sgagias, M. K.; Dixon, K. H.; Anzick, S. L.; Meltzer, P. S.; Cowan, K. H., Isolation of a gene encoding a human reduced folate carrier (RFC1) and analysis of its expression in transport-deficient, methotrexate-resistant human breast cancer cells. *Cancer Res* **1995**, *55* (17), 3790-4.
25. Whetstone, J. R.; Matherly, L. H., The basal promoters for the human reduced folate carrier gene are regulated by a GC-box and a cAMP-response element/AP-1-like element. Basis for tissue-specific gene expression. *J Biol Chem* **2001**, *276* (9), 6350-8.
26. Worm, J.; Kirkin, A. F.; Dzhandzhugazyan, K. N.; Guldberg, P., Methylation-dependent silencing of the reduced folate carrier gene in inherently methotrexate-resistant human breast cancer cells. *J Biol Chem* **2001**, *276* (43), 39990-40000.
27. Yang-Feng, T. L.; Ma, Y. Y.; Liang, R.; Prasad, P. D.; Leibach, F. H.; Ganapathy, V., Assignment of the human folate transporter gene to chromosome 21q22.3 by somatic cell hybrid analysis and in situ hybridization. *Biochem Biophys Res Commun* **1995**, *210* (3), 874-9.
28. Assaraf, Y. G., The role of multidrug resistance efflux transporters in antifolate resistance and folate homeostasis. *Drug Resist Updat* **2006**, *9* (4-5), 227-46.
29. Alt, F. W.; Kellems, R. E.; Bertino, J. R.; Schimke, R. T., Selective multiplication of dihydrofolate reductase genes in methotrexate-resistant variants of cultured murine cells. *J Biol Chem* **1978**, *253* (5), 1357-70.

30. Goldie, J. H.; Krystal, G.; Hartley, D.; Gudauskas, G.; Dedhar, S., A methotrexate insensitive variant of folate reductase present in two lines of methotrexate-resistant L5178Y cells. *Eur J Cancer* **1980**, *16* (12), 1539-46.
31. Then, R. L., Mechanisms of resistance to trimethoprim, the sulfonamides, and trimethoprim-sulfamethoxazole. *Rev Infect Dis* **1982**, *4* (2), 261-9.
32. Smith, H. W., Mutants of *Klebsiella pneumoniae* resistant to several antibiotics. *Nature* **1976**, *259* (5541), 307-8.
33. Smith, D. R.; Calvo, J. M., Nucleotide sequence of dihydrofolate reductase genes from trimethoprim-resistant mutants of *Escherichia coli*. Evidence that dihydrofolate reductase interacts with another essential gene product. *Mol Gen Genet* **1982**, *187* (1), 72-8.
34. Flensburg, J.; Skold, O., Massive overproduction of dihydrofolate reductase in bacteria as a response to the use of trimethoprim. *Eur J Biochem* **1987**, *162* (3), 473-6.
35. Fling, M. E.; Richards, C., The nucleotide sequence of the trimethoprim-resistant dihydrofolate reductase gene harbored by Tn7. *Nucleic Acids Res* **1983**, *11* (15), 5147-58.
36. Smith, D. R.; Calvo, J. M., Nucleotide sequence of the *E coli* gene coding for dihydrofolate reductase. *Nucleic Acids Res* **1980**, *8* (10), 2255-74.
37. Young, H. K.; Skurray, R. A.; Amyes, S. G., Plasmid-mediated trimethoprim-resistance in *Staphylococcus aureus*. Characterization of the first gram-positive plasmid dihydrofolate reductase (type S1). *Biochem J* **1987**, *243* (1), 309-12.
38. Sundstrom, L.; Vinayagamoorthy, T.; Skold, O., Novel type of plasmid-borne resistance to trimethoprim. *Antimicrob Agents Chemother* **1987**, *31* (1), 60-6.
39. [www.uniprot.org](http://www.uniprot.org) (accessed 24 janvier 2012).
40. Howell, E. E.; Booth, C.; Farnum, M.; Kraut, J.; Warren, M. S., A second-site mutation at phenylalanine-137 that increases catalytic efficiency in the mutant aspartate-27----serine *Escherichia coli* dihydrofolate reductase. *Biochemistry* **1990**, *29* (37), 8561-9.
41. Schmitzer, A. R.; Lepine, F.; Pelletier, J. N., Combinatorial exploration of the catalytic site of a drug-resistant dihydrofolate reductase: creating alternative functional configurations. *Protein Eng Des Sel* **2004**, *17* (11), 809-19.

42. Cody, V.; Pace, J.; Makin, J.; Piraino, J.; Queener, S. F.; Rosowsky, A., Correlations of inhibitor kinetics for *Pneumocystis jirovecii* and human dihydrofolate reductase with structural data for human active site mutant enzyme complexes. *Biochemistry* **2009**, *48* (8), 1702-11.
43. Stone, S. R.; Morrison, J. F., Mechanism of inhibition of dihydrofolate reductases from bacterial and vertebrate sources by various classes of folate analogues. *Biochim Biophys Acta* **1986**, *869* (3), 275-85.
44. Amyes, S. G.; Smith, J. T., The purification and properties of the trimethoprim-resistant dihydrofolate reductase mediated by the R-factor, R388. *Eur J Biochem* **1976**, *61* (2), 597-603.
45. Bradrick, T. D.; Beechem, J. M.; Howell, E. E., Unusual binding stoichiometries and cooperativity are observed during binary and ternary complex formation in the single active pore of R67 dihydrofolate reductase, a D2 symmetric protein. *Biochemistry* **1996**, *35* (35), 11414-24.
46. Holland, J. C.; Linn, C. E.; DiGiammarino, E.; Nichols, R.; Howell, E. E., Does R67 dihydrofolate reductase possess a proton donor? *Adv Exp Med Biol* **1993**, *338*, 493-8.
47. Park, H.; Zhuang, P.; Nichols, R.; Howell, E. E., Mechanistic studies of R67 dihydrofolate reductase. Effects of pH and an H62C mutation. *J Biol Chem* **1997**, *272* (4), 2252-8.
48. Krahn, J. M.; Jackson, M. R.; DeRose, E. F.; Howell, E. E.; London, R. E., Crystal structure of a type II dihydrofolate reductase catalytic ternary complex. *Biochemistry* **2007**, *46* (51), 14878-88.
49. Hicks, S. N.; Smiley, R. D.; Stinnett, L. G.; Minor, K. H.; Howell, E. E., Role of Lys-32 residues in R67 dihydrofolate reductase probed by asymmetric mutations. *J Biol Chem* **2004**, *279* (45), 46995-7002.
50. Kamath, G.; Howell, E. E.; Agarwal, P. K., The tail wagging the dog: insights into catalysis in R67 dihydrofolate reductase. *Biochemistry* **2010**, *49* (42), 9078-88.
51. Jackson, M.; Chopra, S.; Smiley, R. D.; Maynard, P. O.; Rosowsky, A.; London, R. E.; Levy, L.; Kalman, T. I.; Howell, E. E., Calorimetric studies of ligand binding in R67 dihydrofolate reductase. *Biochemistry* **2005**, *44* (37), 12420-33.

52. Corbin, E. E.; Prigot, A., Novobiocin; absorption, diffusion, and excretion studies. *Antibiot Annu* **1956**, 392-5.
53. <http://www.hc-sc.gc.ca/>.
54. <http://www.sigmaaldrich.com> (fiche sigalétique du congo red) (accessed 25 janvier 2012).
55. Collins, I.; Workman, P., New approaches to molecular cancer therapeutics. *Nat Chem Biol* **2006**, 2 (12), 689-700.
56. Lindsay, M. A., Target discovery. *Nat Rev Drug Discov* **2003**, 2 (10), 831-8.
57. Horrobin, D. F., Modern biomedical research: an internally self-consistent universe with little contact with medical reality? *Nat Rev Drug Discov* **2003**, 2 (2), 151-4.
58. Butte, A., The use and analysis of microarray data. *Nat Rev Drug Discov* **2002**, 1 (12), 951-60.
59. Phizicky, E.; Bastiaens, P. I.; Zhu, H.; Snyder, M.; Fields, S., Protein analysis on a proteomic scale. *Nature* **2003**, 422 (6928), 208-15.
60. Romero, R.; Kuivaniemi, H.; Tromp, G.; Olson, J., The design, execution, and interpretation of genetic association studies to decipher complex diseases. *Am J Obstet Gynecol* **2002**, 187 (5), 1299-312.
61. Huber, L. A., Is proteomics heading in the wrong direction? *Nat Rev Mol Cell Biol* **2003**, 4 (1), 74-80.
62. Lokey, R. S., Forward chemical genetics: progress and obstacles on the path to a new pharmacopoeia. *Curr Opin Chem Biol* **2003**, 7 (1), 91-6.
63. Matthews, D. J.; Kopczynski, J., Using model-system genetics for drug-based target discovery. *Drug Discov Today* **2001**, 6 (3), 141-149.
64. Fleming, M. P.; Datta, N.; Gruneberg, R. N., Trimethoprim resistance determined by R factors. *Br Med J* **1972**, 1 (5802), 726-8.
65. Amyes, S. G.; Smith, J. T., R-factor trimethoprim resistance mechanism: an insusceptible target site. *Biochem Biophys Res Commun* **1974**, 58 (2), 412-8.
66. Skold, O.; Widh, A., A new dihydrofolate reductase with low trimethoprim sensitivity induced by an R factor mediating high resistance to trimethoprim. *J Biol Chem* **1974**, 249 (13), 4324-5.

67. Gashaw, I.; Ellinghaus, P.; Sommer, A.; Asadullah, K., What makes a good drug target? *Drug Discov Today* **2011**.
68. Pini, A.; Bracci, L., Phage display of antibody fragments. *Curr Protein Pept Sci* **2000**, *1* (2), 155-69.
69. Huovinen, P.; Sundstrom, L.; Swedberg, G.; Skold, O., Trimethoprim and sulfonamide resistance. *Antimicrob Agents Chemother* **1995**, *39* (2), 279-89.
70. Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T., Determinants of aminoglycoside-binding specificity for rRNA by using mass spectrometry. *Proc Natl Acad Sci U S A* **1999**, *96* (18), 10129-33.
71. Swayze, E. E.; Jefferson, E. A.; Sannes-Lowery, K. A.; Blyn, L. B.; Risen, L. M.; Arakawa, S.; Osgood, S. A.; Hofstadler, S. A.; Griffey, R. H., SAR by MS: a ligand based technique for drug lead discovery against structured RNA targets. *J Med Chem* **2002**, *45* (18), 3816-9.
72. Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W., Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **1996**, *274* (5292), 1531-4.
73. Jhoti, H.; Cleasby, A.; Verdonk, M.; Williams, G., Fragment-based screening using X-ray crystallography and NMR spectroscopy. *Curr Opin Chem Biol* **2007**, *11* (5), 485-93.
74. Ward, W. H.; Holdgate, G. A., Isothermal titration calorimetry in drug discovery. *Prog Med Chem* **2001**, *38*, 309-76.
75. Perspicace, S.; Banner, D.; Benz, J.; Muller, F.; Schlatter, D.; Huber, W., Fragment-based screening using surface plasmon resonance technology. *J Biomol Screen* **2009**, *14* (4), 337-49.
76. Nienaber, V. L.; Richardson, P. L.; Klighofer, V.; Bouska, J. J.; Giranda, V. L.; Greer, J., Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nat Biotechnol* **2000**, *18* (10), 1105-8.
77. Hartshorn, M. J.; Murray, C. W.; Cleasby, A.; Frederickson, M.; Tickle, I. J.; Jhoti, H., Fragment-based lead discovery using X-ray crystallography. *J Med Chem* **2005**, *48* (2), 403-13.



78. Zhang, J. H.; Chung, T. D.; Oldenburg, K. R., A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J Biomol Screen* **1999**, *4* (2), 67-73.
79. Schuffenhauer, A.; Ruedisser, S.; Marzinzik, A. L.; Jahnke, W.; Blommers, M.; Selzer, P.; Jacoby, E., Library design for fragment based screening. *Curr Top Med Chem* **2005**, *5* (8), 751-62.
80. Gribbon, P.; Sewing, A., High-throughput drug discovery: what can we expect from HTS? *Drug Discov Today* **2005**, *10* (1), 17-22.
81. Fink, T.; Reymond, J. L., Virtual exploration of the chemical universe up to 11 atoms of C, N, O, F: assembly of 26.4 million structures (110.9 million stereoisomers) and analysis for new ring systems, stereochemistry, physicochemical properties, compound classes, and drug discovery. *J Chem Inf Model* **2007**, *47* (2), 342-53.
82. Bembenek, S. D.; Tounge, B. A.; Reynolds, C. H., Ligand efficiency and fragment-based drug discovery. *Drug Discov Today* **2009**, *14* (5-6), 278-83.
83. Murray, C. W.; Rees, D. C., The rise of fragment-based drug discovery. *Nat Chem* **2009**, *1* (3), 187-92.
84. Ripphausen, P.; Nisius, B.; Peltason, L.; Bajorath, J., Quo vadis, virtual screening? A comprehensive survey of prospective applications. *J Med Chem* **2010**, *53* (24), 8461-7.
85. [www.zinc.docking.org](http://www.zinc.docking.org) (accessed 20 mars 2012).
86. Irwin, J. J.; Shoichet, B. K., ZINC--a free database of commercially available compounds for virtual screening. *J Chem Inf Model* **2005**, *45* (1), 177-82.
87. Geppert, H.; Vogt, M.; Bajorath, J., Current Trends in Ligand-Based Virtual Screening: Molecular Representations, Data Mining Methods, New Application Areas, and Performance Evaluation. *Journal of Chemical Information and Modeling* **2010**, *50* (2), 205-216.
88. Kerns, E. H., High throughput physicochemical profiling for drug discovery. *J Pharm Sci* **2001**, *90* (11), 1838-58.
89. Di, L.; Kerns, E. H., Profiling drug-like properties in discovery research. *Curr Opin Chem Biol* **2003**, *7* (3), 402-8.

90. McGovern, S. L.; Caselli, E.; Grigorieff, N.; Shoichet, B. K., A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J Med Chem* **2002**, *45* (8), 1712-22.
91. Venkatesh, S.; Lipper, R. A., Role of the development scientist in compound lead selection and optimization. *J Pharm Sci* **2000**, *89* (2), 145-54.
92. Yan, Z.; Caldwell, G. W., Metabolism profiling, and cytochrome P450 inhibition & induction in drug discovery. *Curr Top Med Chem* **2001**, *1* (5), 403-25.
93. Hefti, F. F., Requirements for a lead compound to become a clinical candidate. *BMC Neurosci* **2008**, *9 Suppl 3*, S7.
94. Redfern, W. S.; Carlsson, L.; Davis, A. S.; Lynch, W. G.; MacKenzie, I.; Palethorpe, S.; Siegl, P. K.; Strang, I.; Sullivan, A. T.; Wallis, R.; Camm, A. J.; Hammond, T. G., Relationships between preclinical cardiac electrophysiology, clinical QT interval prolongation and torsade de pointes for a broad range of drugs: evidence for a provisional safety margin in drug development. *Cardiovasc Res* **2003**, *58* (1), 32-45.
95. Schulz, M. N.; Hubbard, R. E., Recent progress in fragment-based lead discovery. *Curr Opin Pharmacol* **2009**, *9* (5), 615-21.
96. Jencks, W. P., On the attribution and additivity of binding energies. *Proc Natl Acad Sci U S A* **1981**, *78* (7), 4046-50.
97. Murray, C. W.; Verdonk, M. L., The consequences of translational and rotational entropy lost by small molecules on binding to proteins. *J Comput Aided Mol Des* **2002**, *16* (10), 741-53.
98. Congreve, M.; Chessari, G.; Tisi, D.; Woodhead, A. J., Recent developments in fragment-based drug discovery. *J Med Chem* **2008**, *51* (13), 3661-80.
99. Geschwindner, S.; Olsson, L. L.; Albert, J. S.; Deinum, J.; Edwards, P. D.; de Beer, T.; Folmer, R. H., Discovery of a novel warhead against beta-secretase through fragment-based lead generation. *J Med Chem* **2007**, *50* (24), 5903-11.
100. Frederickson, M.; Callaghan, O.; Chessari, G.; Congreve, M.; Cowan, S. R.; Matthews, J. E.; McMenamin, R.; Smith, D. M.; Vinkovic, M.; Wallis, N. G., Fragment-based discovery of mexiletine derivatives as orally bioavailable inhibitors of urokinase-type plasminogen activator. *J Med Chem* **2008**, *51* (2), 183-6.

101. Card, G. L.; Blasdel, L.; England, B. P.; Zhang, C.; Suzuki, Y.; Gillette, S.; Fong, D.; Ibrahim, P. N.; Artis, D. R.; Bollag, G.; Milburn, M. V.; Kim, S. H.; Schlessinger, J.; Zhang, K. Y., A family of phosphodiesterase inhibitors discovered by cocrystallography and scaffold-based drug design. *Nat Biotechnol* **2005**, *23* (2), 201-7.
102. Oltersdorf, T.; Elmore, S. W.; Shoemaker, A. R.; Armstrong, R. C.; Augeri, D. J.; Belli, B. A.; Bruncko, M.; Deckwerth, T. L.; Dinges, J.; Hajduk, P. J.; Joseph, M. K.; Kitada, S.; Korsmeyer, S. J.; Kunzer, A. R.; Letai, A.; Li, C.; Mitten, M. J.; Nettesheim, D. G.; Ng, S.; Nimmer, P. M.; O'Connor, J. M.; Oleksijew, A.; Petros, A. M.; Reed, J. C.; Shen, W.; Tahir, S. K.; Thompson, C. B.; Tomaselli, K. J.; Wang, B.; Wendt, M. D.; Zhang, H.; Fesik, S. W.; Rosenberg, S. H., An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* **2005**, *435* (7042), 677-81.
103. Sharpless, K. B. M., R., In situ click chemistry: a powerful means for lead discovery. *Expert Opin. Drug DiscoVery* **2007**, *1*, 525-538.
104. Mocharla, V. P.; Colasson, B.; Lee, L. V.; Roper, S.; Sharpless, K. B.; Wong, C. H.; Kolb, H. C., In situ click chemistry: enzyme-generated inhibitors of carbonic anhydrase II. *Angew Chem Int Ed Engl* **2004**, *44* (1), 116-20.
105. Shoichet, B. K., Virtual screening of chemical libraries. *Nature* **2004**, *432* (7019), 862-5.
106. Yuriev, E.; Agostino, M.; Ramsland, P. A., Challenges and advances in computational docking: 2009 in review. *J Mol Recognit* **2011**, *24* (2), 149-64.
107. MolegroVirtualDocker, User manual **2011**, 273.
108. Huang, N.; Kalyanaraman, C.; Irwin, J. J.; Jacobson, M. P., Physics-based scoring of protein-ligand complexes: enrichment of known inhibitors in large-scale virtual screening. *J Chem Inf Model* **2006**, *46* (1), 243-53.
109. Huang, S. Y.; Grinter, S. Z.; Zou, X., Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. *Phys Chem Chem Phys* **2010**, *12* (40), 12899-908.
110. Allen, F. H., The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr B* **2002**, *58* (Pt 3 Pt 1), 380-8.
111. Hartman, G. D.; Egbertson, M. S.; Halczenko, W.; Laswell, W. L.; Duggan, M. E.; Smith, R. L.; Naylor, A. M.; Manno, P. D.; Lynch, R. J.; Zhang, G.; et al., Non-

peptide fibrinogen receptor antagonists. 1. Discovery and design of exosite inhibitors. *J Med Chem* **1992**, 35 (24), 4640-2.

112. Leban, J.; Saeb, W.; Garcia, G.; Baumgartner, R.; Kramer, B., Discovery of a novel series of DHODH inhibitors by a docking procedure and QSAR refinement. *Bioorg Med Chem Lett* **2004**, 14 (1), 55-8.

113. Clark, D. E., What has virtual screening ever done for drug discovery? *Expert Opin. Drug Discov* **2008**, 3 (8), 841-851.

114. Cheng, Y.; Prusoff, W. H., Relationship between the inhibition constant ( $K_i$ ) and the concentration of inhibitor which causes 50 per cent inhibition ( $I_{50}$ ) of an enzymatic reaction. *Biochem Pharmacol* **1973**, 22 (23), 3099-108.

115. Dixon, M., The determination of enzyme inhibitor constants. *Biochem J* **1953**, 55 (1), 170-1.

## Chapitre 2 - Design de l'essai automatisé

À la section 1.1.2, on a vu plusieurs possibilités de détecter la liaison entre une protéine et des ligands. On a vu que le test doit être automatisable et rapide afin d'être en mesure de cribler un grand nombre de composés. Par la suite, on a vu qu'un test d'activité doit être développé afin de déterminer si les composés qui se lient à la cible provoquent la réponse biologique recherchée. Dans le cas de la DHFR R67, le test d'activité était facilement adaptable à un criblage automatisé, ainsi, on a été en mesure de chercher directement des inhibiteurs de la DHFR R67.

### Section 2.1 - Mesure de l'activité de la DHFR R67

L'activité de la DHFR R67 peut être suivie en temps réel par spectrophotométrie, en suivant l'absorbance du milieu réactionnel. Selon la loi de Beer-Lambert (équation 8), l'absorbance est directement reliée à la concentration d'un analyte.

$$A = \epsilon lc \quad (8)$$

A = absorbance

$\epsilon$  = coefficient d'absorptivité molaire

l = longueur du chemin optique

c = concentration de l'analyte

On a vu que la DHFR R67 catalyse la réduction du dihydrofolate (DHF) en tétrahydrofolate (THF) par l'oxydation du cofacteur nicotinamide adénine dinucléotide phosphate (NADPH). Les réactifs, le DHF et le NADPH absorbent fortement à 340 nm tandis que les produits, le THF et le  $\text{NADP}^+$ , absorbent faiblement à cette longueur d'onde. Ainsi, l'absorbance du milieu réactionnel diminue au fil de la réaction ce qui permet de déterminer l'activité de la DHFR R67.

Étant donné qu'on recherche initialement des inhibiteurs faibles, on utilise des concentrations de DHF et NADPH suffisamment faibles pour ne pas masquer une faible

inhibition compétitive tout en donnant lieu à un changement d'absorbance significatif (signal). Il fut déterminé qu'une concentration de 50  $\mu\text{M}$  de DHF et de NADPH était un bon compromis pour donner lieu à un signal linéaire sur plusieurs minutes permettant une détermination robuste de la vitesse initiale en présence de composés potentiellement inhibiteurs.

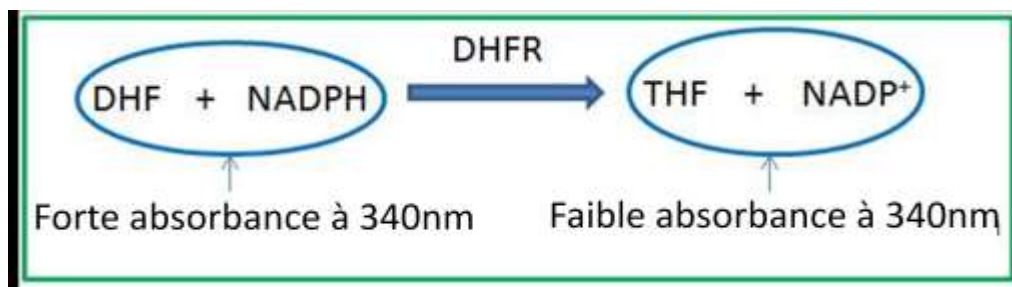


Figure 2.1 : Mesure de l'activité de la DHFR R67 par spectrophotométrie.  $\Delta\epsilon$  réactifs/produits =  $12.8 \text{ mM}^{-1}\text{cm}^{-1}$

Cet essai est facilement automatisable, car il peut être réalisé dans des plaques de 96 puits. Avec un volume par puits de 100-200  $\mu\text{L}$ , les essais ne nécessitent que peu de matériel (DHF, NADPH, enzyme et composé). À l'aide d'une station robotisée de pipetage Biomeck NX1 (BeckmanCoulter) jumelée à une lectrice à plaques DTX880 (BeckmanCoulter), les échantillons peuvent être rapidement préparés et analysés.



Figure 2.2 : Station robotisée de pipetage utilisée lors du criblage d'inhibiteur contre la DHFR R67.

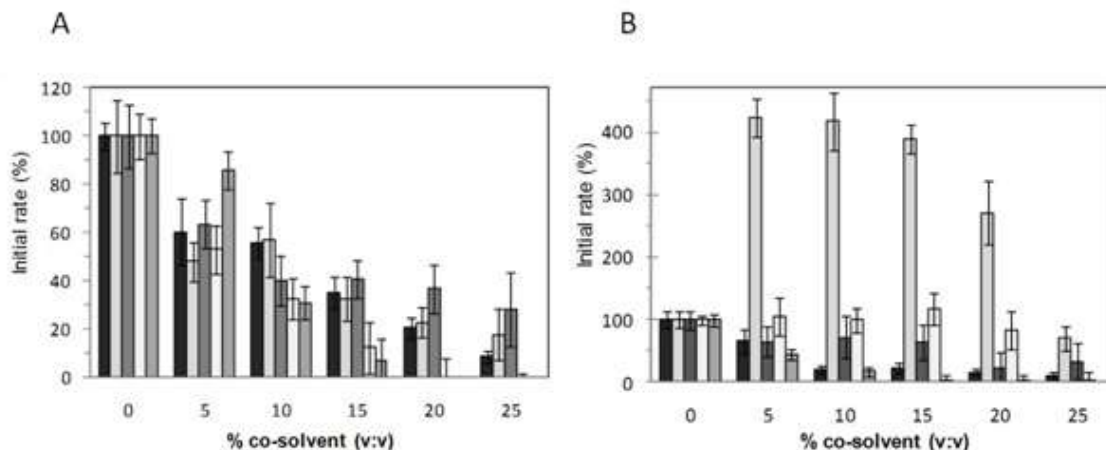
## **Section 2.2 - Détermination des conditions de criblage**

Le test d'activité de la DHFR R67 présenté à la section 2.1 à plusieurs limitations comme la solubilité des composés en milieu aqueux, l'absorbance des composés et les variations de pH dues aux composés. Une mauvaise solubilisation des composés peut rendre le test impossible, car la DHFR R67 ne sera pas en présence d'une concentration adéquate de composés. S'il y a lieu, un précipité peut causer une interférence avec le signal et/ou former des agrégats qui pourraient conduire à de faux positifs. Si un composé absorbe fortement à 340 nm, la longueur d'onde à laquelle les mesures sont prises, l'absorbance du milieu réactionnel sera élevée, pouvant dépasser la limite de détection de l'appareil et mener à de faux positifs. Finalement, on a vu que la DHFR R67 est très sensible au pH. Une augmentation de pH aussi faible que 0.1 unité peut faire diminuer l'activité de la DHFR R67 d'environ 10 %, ce qui peut mener à de faux positifs (voir chapitre 1, section 1.2).

### **Section 2.2.1 - Solubilité des composés**

Afin d'améliorer la solubilité de composé hydrophobes, nous avons utilisé un système de co-solvant, c'est-à-dire, on ajoute un solvant organique au milieu réactionnel. Le désavantage d'utiliser des solvants organiques dans un milieu réactionnel enzymatique est qu'ils nuisent généralement aux enzymes. Ainsi l'ajout de solvant augmente la solubilité des composés hydrophobes, mais diminue l'activité enzymatique. Ainsi, il faut trouver un compromis entre la solubilité des composés et l'activité enzymatique.

Afin de déterminer un co-solvant organique et une concentration à utiliser, on a testé l'activité de la DHFR R67 et de la DHFR humaine dans le méthanol (MeOH), l'acétonitrile (ACN), le tétrahydrofurane (THF, à ne pas confondre avec le tétrahydrofolate), le diméthylformamide (DMF) et le diméthylsulfoxyde (DMSO) pour des concentrations de 0 % à 25 %.



**Figure 2.3 : Activité des DHFR en présence de différentes concentrations de co-solvants organiques.** A) DHFR R67. B) DHFR humaine. ( ■ ) MeOH, ( □ ) ACN, ( ■ ) THF, ( □ ) DMF, ( ■ ) DMSO. Les valeurs correspondent à la moyenne  $\pm$  écart-type de 3 expériences indépendantes fait chacune en triplicate. Ces résultats, que j'ai obtenus durant l'automne 2010, figurent également dans le mémoire de Natalia Kadnikova, étudiante à la maîtrise au laboratoire du Dr. Pelletier.

Dans le cas de la DHFR R67 (figure 2.3 A), à 5 % de co-solvant, le DMSO est le solvant qui perturbe le moins l'activité de la DHFR R67, environ de 80% de l'activité, alors que 5% des autres co-solvant réduit l'activité d'environ 50%. Lorsqu'on augmente la concentration de co-solvant à 10 %, l'activité de la DHFR R67 dans le THF, le DMF et le DMSO diminue à 30 % tandis que l'activité dans le MeOH et l'ACN demeure aux environs de 50 %.

Dans le cas de la DHFR humaine (figure 2.3 B), 5% de co-solvant ne permet que de conserver qu'environ 50% de l'activité de la DHFR humaine pour le MeOH, THF et DMSO. À 10 % de ceux-ci, l'activité diminue à environ 20 %. Toutefois, l'activité de la DHFR humaine a augmenté de 3 à 4 fois, jusqu'à une concentration de 20 % d'ACN, et le DMF n'a pas affecté l'activité jusqu'à une concentration de 20 %.

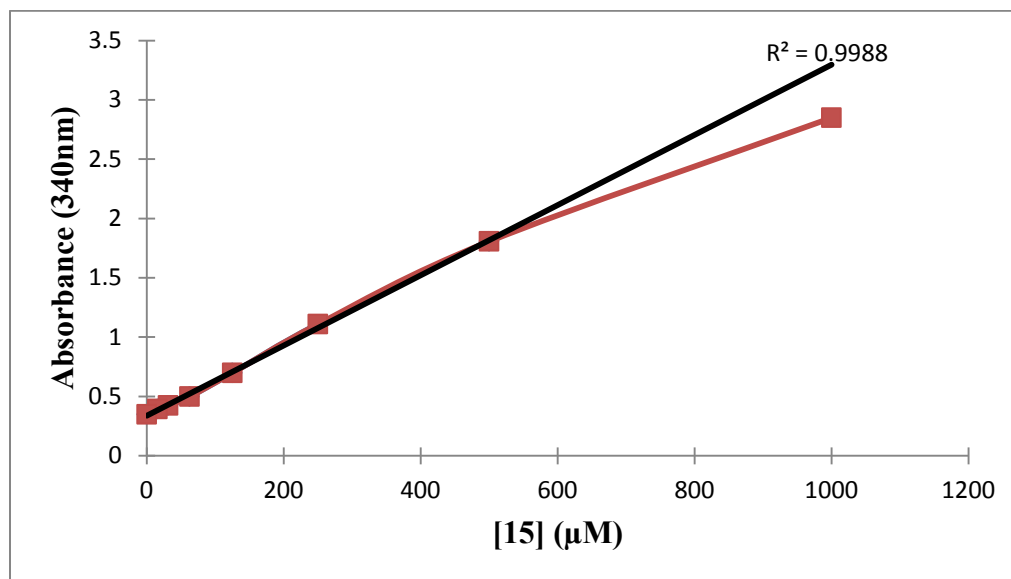
Au niveau de la rétention de l'activité pour les deux DHFR, l'ACN serait le meilleur co-solvant. Par contre, afin de développer une méthode de criblage d'inhibiteurs, il faut aussi déterminer la capacité des co-solvants à dissoudre des composés. Le DMSO a généralement démontré un meilleur pouvoir de solubilisation des composés hydrophobes. Ainsi, le DMSO semble être le meilleur compromis entre la



rétenction de l'activité des DHFR et la solubilité des composés, donc on a choisi d'utiliser une concentration de 10 % DMSO pour effectuer le criblage d'inhibiteurs potentiels envers la DHFR R67, avec la DHFR humaine comme contrôle.

## Section 2.2.2 - Absorbance des composés

Lors des essais d'une DHFR, la mesure de l'absorbance du milieu réactionnel se fait à 340 nm puisque c'est à cette longueur d'onde que la différence d'absorbance entre les réactifs et les produits est maximale. Ainsi, des composés testés qui absorbent fortement à 340 nm peuvent causer de l'interférence avec le signal de l'activité enzymatique, faisant dépasser la limite de détection de la lectrice à plaques. La limite de détection de l'appareil utilisé, un DTX880 de BeckmanCoulter, fut déterminée en utilisant un composé qui absorbe fortement à 340 nm (figure 2.4). Il y a une linéarité du signal jusqu'à une absorbance d'environ 2, puis on sort de la zone de linéarité de l'appareil. Ainsi, il faut rejeter toute mesure qui aurait une absorbance au-dessus de 2.



**Figure 2.4: Limite de détection de la lectrice à plaque DTX880.** L'analyte examiné est le composé 15 (chapitre 4). En rouge, on peut voir l'absorbance mesurée pour des concentrations de 0 μM, 16 μM, 32 μM, 64 μM, 125 μM, 250 μM, 500 μM, 1000 μM. En noir, il s'agit de la régression linéaire pour tout les points ayant une absorbance au dessous de 2. Les mesures ont été faites en triplicata lors d'une unique expérience  $R^2 = 0,9988$ .

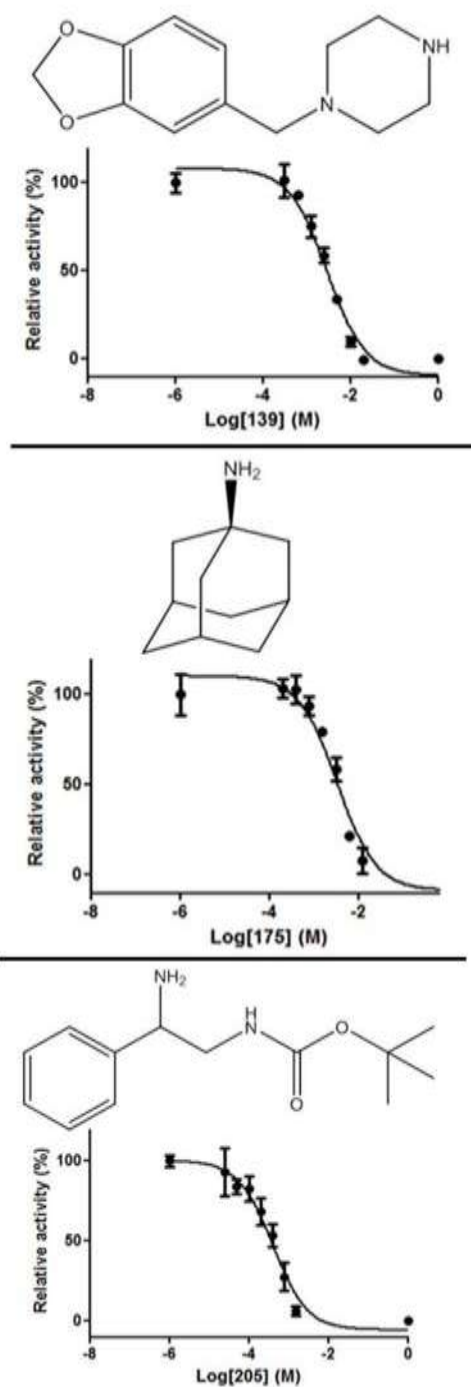
Dans le but de pouvoir étudier les composés qui absorbent fortement à 340 nm, on peut diminuer le chemin optique en diminuant le volume. En effet, un des avantages d'utiliser une lectrice à plaques est que le faisceau lumineux est perpendiculaire à l'échantillon. Ainsi, si on réduit le volume de l'échantillon, on réduit le chemin optique, ce qui réduit l'absorbance (équation 8). En diminuant le volume de 200  $\mu$ L à 100  $\mu$ L, l'absorbance diminue d'un facteur d'environ 2 (tableau 2.1). Cette approche permet de tester des concentrations de composés qui auraient dépassé la limite d'absorbance de 2 pour un volume de 200  $\mu$ L. Toutefois, les plus faible volumes peuvent augmenter l'erreur expérimentale.

**Tableau 2.1 : Absorbance pour des volume de 100  $\mu$ L et 200  $\mu$ L du composé 15 (chapitre 4)**

Concentration du 15 ( $\mu$ M)	Absorbance à 100 $\mu$ L	Absorbance à 200 $\mu$ L
250	0.83	1.53
125	0.41	0.72
64	0.21	0.33

### Section 2.2.3 - pH du milieu réactionnel

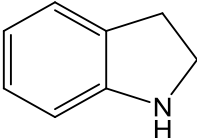
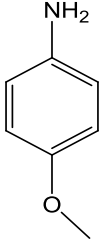
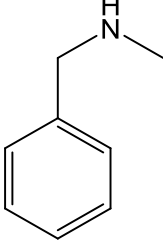
Étant donné qu'on identifiera en premier lieu des inhibiteurs faibles (ordre du millimolaire), on teste les composés à de hautes concentrations (millimolaire), ce qui peut modifier le pH du milieu réactionnel s'il n'est pas fortement tamponné. La figure 2.5, illustre des  $IC_{50}$  déterminés à partir de composés qui n'étaient pas des inhibiteurs réels de la DHFR R67 (figure 2.5). Ces trois composés sont des bases faibles dû à la présence d'amine (le tampon utilisé dans ces exemple est 50mM de phosphate de potassium). Dû aux grandes concentrations de composés testés, le pH a augmenté légèrement diminuant l'activité de la DHFR R67 qui est sensible au pH (figure 1.7).



**Figure 2.5 : Titrage acido-basique de l'activité de la DHFR R67 par des bases faibles.** Les numéros correspondent aux numéros attribués dans la banque de fragment. Composé 139 - Pipéronylpipérazine, 175 - Amantadine, 205 -Ester tert-butyl d'acide (2-Amino-2-phényl-éthyl)-carbamique.

Étant donné que 50 mM de phosphate de potassium ne suffisait pas à tamponner adéquatement le système, nous avons testé l'HEPES en forte concentration. L'HEPES fut choisi car sa zone tampon est de 6.8 à 8: ce qui correspond à la zone d'activité de la DHFR R67. Le pH déterminé pour la prise des mesures est 7, ce qui correspond au pH physiologique. Le tableau 2.2, rapporte le pH après l'ajout de 10 mM de composés typiques choisis dans le criblage à 200 mM et 400 mM d'HEPES. L'HEPES à 200mM n'était pas suffisant pour garder un pH stable à la valeur de 7 à l'ajout de 10 mM de ces composés. Par contre, 400 mM d'HEPES était suffisant pour garder le pH stable à 7. Le pH est également resté stable à 7 lors d'ajout d'acide fort (10 mM HCl) ou de base forte (10mM NaOH), ce qui démontre que des bases faibles et des acides faible ne pourraient modifier le pH (à concentration similaire). Les valeurs de pH rapportées sont approximatives puisqu'elles ont été prises avec du papier pH. Ainsi, on a choisi le tampon HEPES à 400mM pour effectuer le criblages de fragments.

**Tableau 2.2 : pH de la solution après l'ajout de 10 mM des composés types criblés\***

Composé*	Structure	HEPES 200 mM	HEPES 400 mM
17		8	7
27		8	7
33		8	7

\*Les numéros correspondent aux numéros attribués dans la banque de fragments.

17 - Indoline, 27 - Anisidine, 33 - N-méthylbenzylamine

## **Section 2.3 - Fiabilité de l'essai**

Ayant déterminé les conditions de criblage et automatisé l'essai, on a déterminé le Zscore de l'essai afin de confirmer qu'il était fiable et qu'il pouvait être utilisé pour le criblage de composés sur la DHFR R67. Afin de déterminer le Zscore, 48 contrôles positifs (avec enzyme) et 48 contrôles négatifs (sans enzymes) ont été faits en une plaque de 96 puits. Le Zscore de l'essai est de 0.8, ce qui est nettement au-dessus de la valeur de 0.5, donc l'essai est fiable. Ainsi, le criblage de composé a été entrepris.

## **Chapitre 3 - Design à base de fragments**

**Article 1. Design à base de fragments de bis-benzimidazoles symétriques comme inhibiteurs sélectifs de l'enzyme dihydrofolate réductase de type II R67 résistante au triméthoprim.**

**Article 1. Fragment-based design of symmetrical bis-benzimidazoles as selective inhibitors of the trimethoprim-resistant, Type II R67 dihydrofolate reductase**

## Section 3.1 - Présentation de l'article et mise en contexte

Le triméthoprim (TMP) est un inhibiteur de la DHFR chromosomale bactérienne. Cette enzyme est responsable de la catalyse du dihydrofolate (DHF) en tétrahydrofolate (THF) chez les bactéries, qui lui, est essentiel à la synthèse des purines. De ce fait, le TMP empêche la prolifération bactérienne. Le TMP est le principal antibiotique utilisé pour soigner les infections urinaires chez l'humain. Il est aussi utilisé dans l'aquaculture et l'élevage de bovins. La grande utilisation du triméthoprim a mené à l'apparition de divers mécanismes de résistance dont l'expression d'une enzyme plasmidique: la DHFR R67. La DHFR R67 n'est pas affectée par le TMP aux concentrations utiles, et peut ainsi remplacer la DHFR chromosomale lorsque celle-ci est inhibée par le TMP. Elle permet ainsi la prolifération bactérienne en présence de TMP.

Dans le but de briser la résistance au triméthoprim conféré par la DHFR R67, nous cherchons à développer des inhibiteurs spécifiques à cette enzyme. Il existe présentement deux inhibiteurs non-spécifiques de la DHFR R67, soit la novobiocine et le rouge de congo. Dû à la faible puissance de la novobiocine et de la toxicité du rouge de congo, il est essentiel de développer des inhibiteurs de la DHFR R67 *de novo*.

Le design à base de fragments est de plus en plus utilisé de nos jours et plusieurs candidats-médicaments ont été développés selon cette approche. Cette approche fut utilisée lors de l'étude présentée dans ce chapitre. Le criblage d'une librairie d'environ 100 fragments a permis d'identifier 7 inhibiteurs faibles de la DHFR R67 (ordre du mM). En construisant des molécules plus complexes à partir de ces 7 fragments, il a été possible d'identifier deux molécules plus puissantes qui inhibent sélectivement la DHFR R67 dans l'ordre du micromolaire. L'étude d'analogues de ces inhibiteurs a permis de démontrer que le groupement benzimidazole et la fonction acide carboxylique étaient importants dans la liaison des inhibiteurs micromolaires.

Le profilage de ces inhibiteurs a permis de démontrer que ce sont des inhibiteurs compétitifs envers le DHF et le NADPH, qu'au moins deux molécules d'inhibiteurs se

lient dans le site actif de l'enzyme et qu'ils ne sont pas hautement toxiques envers les fibroblastes 3T6. Finalement, des études d'arrimage moléculaire ont permis la conception de modèles qui peuvent expliquer les interactions impliquées dans la liaison de ces inhibiteurs. Ces modèles pourraient aussi expliquer les différentes affinités observées entre les inhibiteurs et leurs analogues lors des tests *in vitro*.

Ma participation à ce projet comprend la mise au point des conditions de criblage (chapitre 2). Elle s'est poursuivie lors du criblage et de l'identification, de l'évaluation de la puissance et de la sélectivité des inhibiteurs faibles. Par la suite, j'ai entrepris l'identification et l'évaluation de la puissance et de la sélectivité des inhibiteurs micromolaires ainsi que la détermination de leur modes d'inhibition et la stœchiométrie. Finalement, j'ai fait les études d'arrimage moléculaire. Maximilian Ebert a contribué à la détermination de la stoechiométrie des inhibiteurs. Jacynthe Toulouse a participé à l'identification des inhibiteurs faibles durant son stage d'été sous ma supervision. Delphine Forge a fourni les résultats de cytotoxicité. Toutes les synthèses ont été réalisées par Delphine Forge, Florent Perron, Annie Mayence et Tien L. Huang du laboratoire du Dr. Vanden Eyde. Natalia Kadnikova a offert une contribution théorique et des résultats préliminaires lors de la mise au point des conditions de criblages.

J'ai écrit toute la publication sous la supervision du Dr. Joelle Pelletier excepté la partie stœchiométrie qui a été écrite par Maximilian Hebert, toujours sous la supervision du Dr. Joelle Pelletier.



**Fragment-based design of symmetrical bis-benzimidazoles  
as selective inhibitors of the trimethoprim-resistant, Type II  
R67 dihydrofolate reductase**

Dominic Bastien<sup>1</sup>, Maximilian C. C. J. C. Ebert<sup>1</sup>, Delphine Forge<sup>2</sup>, Jacynthe  
Toulouse<sup>1</sup>, Natalia Kadnikova<sup>3</sup>, Florent Perron<sup>2</sup>, Annie Mayence<sup>4</sup>, Tien L. Huang<sup>4</sup>, Jean  
Jacques Vanden Eynde<sup>2</sup> and Joelle N. Pelletier<sup>\*,1,3</sup>

From <sup>1</sup> Département de biochimie, Université de Montréal, <sup>2</sup> Laboratoire de chimie  
organique, Université de Mons, <sup>3</sup> Département de chimie, Université de Montréal, and  
<sup>4</sup> Division of Basic Pharmaceutical Sciences, Xavier University of Louisiana.

*J. Med. Chem.*, 2012, **55**, 3182-3192

## Abstract

The continuously increasing use of trimethoprim as a common antibiotic for medical use and for prophylactic application in terrestrial and aquatic animal farming has increased its prevalence in the environment. This has been accompanied by increased drug resistance, generally in the form of alterations in the drug target, dihydrofolate reductase (DHFR). The most highly resistant variants of DHFR are known as Type II DHFR, among which R67 DHFR is the most broadly studied variant. We report the first attempt at designing specific inhibitors to this emerging drug target by fragment-based design. Detection of inhibition in R67 DHFR was accompanied by parallel monitoring of the human DHFR, as an assessment of compound selectivity. By those means, small aromatic molecules of 150 to 250 g/mole (fragments) inhibiting R67 DHFR selectively in the low millimolar range were identified. More complex, symmetrical bis-benzimidazoles and a bis-carboxyphenyl were then assayed as fragment-based leads, which procured selective inhibition of the target in the low micromolar range ( $K_i = 2 - 4 \mu\text{M}$ ). The putative mode of inhibition is discussed according to molecular modelling supported by *in vitro* tests.

## Introduction

Trimethoprim (TMP) has been used as an antibiotic worldwide for the past 50 years, frequently in combination with sulfamethoxazole<sup>1</sup>. Today, TMP remains a primary treatment of urinary *E. coli* infections, and is used to treat respiratory infections such as pneumonia. It has been used to treat sexually transmitted infections such as gonorrhoea but this has resulted in high TMP resistance<sup>2</sup>. As well, TMP with sulfamethoxazole shows potential for treatment of tuberculosis<sup>2b</sup>, which causes an estimated 1.7 million deaths each year, and was recommended by the WHO in 2011 to reduce mortality due to opportunistic infections in adults living with HIV/AIDS<sup>3</sup>. In addition to medical use, TMP is also used for veterinary applications, including prophylactic use in porcine, bovine<sup>4</sup>, salmon<sup>5</sup> and shrimp farming<sup>5-6</sup>. These result in large-scale environmental dissemination of the antibiotic<sup>7</sup>, which is inevitably accompanied by horizontal transfer of plasmids carrying multiple drug resistance genes to other microbes, including human pathogens<sup>5, 8</sup>. As a result, there is a growing concern world-wide over the increased incidence in TMP resistance and its impact on the long-term effectiveness of this low-cost antibiotic.

TMP is a specific inhibitor of bacterial dihydrofolate reductases (DHFR). DHFRs are oxidoreductases that catalyze the reduction of dihydrofolate (DHF) to tetrahydrofolate, using nicotinamide dinucleotide phosphate (NADPH) as a hydride-donating cofactor. DHFR is an essential enzyme for synthesis of DNA precursors (purines, thymidylate) in all living cells, where it is chromosomally encoded and has been highly conserved throughout evolution<sup>9</sup>. Because of its essential role in cellular proliferation, DHFR is a central target in the control of proliferative diseases, including microbial infections. There are sufficient structural differences between human DHFR (hDHFR) and its microbial counterparts to allow for selective inhibition by compounds such as pyrimethamine and cycloguanil, with high clinical efficacy<sup>10</sup>. Trimethoprim is such a selective inhibitor of microbial DHFRs; it binds to *Escherichia coli* DHFR approximately 4 orders of magnitude more tightly than it binds to hDHFR<sup>11</sup>. Among the mechanisms of trimethoprim resistance that have been reported<sup>12</sup>, R-plasmid encoded type II DHFR<sup>13</sup> constitutes an important health concern in both humans and livestock by

sustaining bacterial proliferation<sup>14</sup>. Type II R67 DHFR catalyzes the same reaction as chromosomally encoded DHFR<sup>15</sup> albeit with a lower efficiency, yet it evades the action of TMP, providing bacterial drug resistance<sup>13, 16</sup>. Indeed, R67 DHFR is highly resistant to TMP, with  $IC_{50}$  and  $K_i > 10^6$  fold higher than its chromosomal counterpart<sup>16-17</sup>. In contrast to the DHF substrate, TMP has no amide with which to H-bond with the Ile68 backbone. In addition, R67 DHFR has no carboxylate in its hydrophobic active site<sup>18</sup>, thus cannot form electrostatic interactions with the protonated TMP as seen in the chromosomal DHFR<sup>19</sup>. Indeed, R67 DHFR is genetically and structurally unrelated to chromosomal DHFRs, resulting in important structural differences which form the basis of its intrinsic trimethoprim resistance. This enzyme is a homotetramer, where four identical protomers form a symmetrical active site pore that binds a single DHF/NADPH pair for turnover<sup>19</sup>. The active site cavity is a ‘binding hot-spot’ of relatively low specificity that can bind two DHF or two NADPH molecules at once<sup>17, 19</sup> and reduce some folate analogs<sup>15</sup>. This binding promiscuity runs counter the excruciatingly high binding stringency of hDHFR, which should allow for discovery of selective inhibitors of R67 DHFR that will not significantly inhibit hDHFR. Taken along with TMP, inhibitors of R67 DHFR would break TMP resistance granted by R67 DHFR.

Novobiocin, a DNA gyrase inhibitor, inhibits the activity of R67 DHFR as does congo red, a diazo dye<sup>17</sup>. Neither is specific to R67 DHFR, yet they provide structural information relative to functional groups of interest in inhibitor design. In this work, we applied fragment-based inhibitor screening<sup>20</sup> to identify simple molecules that inhibit R67 selectively. While such small molecules may exhibit only a weak inhibitory effect due to their small size, they may form high quality interactions as expressed by the ratio of affinity to number of heavy atoms, or ‘ligand efficiency’ (LE)<sup>21</sup>. Inhibitory ‘fragments’ allow the subsequent identification or design of more complex molecules sharing those substructures, which may display improved inhibition or other properties relevant to drug design. Molecules resulting from fragment-based design may require less optimization than their counterparts identified by more conventional high-throughput screening (HTS) campaigns, which may contain non-desirable groups<sup>21</sup>.

We seek molecules that bind within the active site in an effort to increase likelihood of selectivity. Thus, screening for inhibition was carried out against a panel of approximately 100 simple compounds (or fragments) with a molecular weight of 150-250g/mol. The fragments screened generally comprise cycles – often aromatic – and were generally of overall elongated structure. Nitrogen and oxygen were preferred heteroatoms. While broader chemical diversity was also included, the above features were prioritized as they imitate features of the nicotinamide ring of NADPH and of the pteroyl ring of DHF which are involved in binding to the centre of the active site cavity<sup>19</sup>. In order to restrict downstream efforts to molecules showing the best potential for selective inhibitor development, we screened the target bacterial R67 DHFR in parallel with human DHFR, which should be spared. Here, we report seven fragments which inhibit the R67 DHFR with an IC<sub>50</sub> in the high micromolar to low millimolar range. Symmetrical bis-benzimidazole and bis-carboxyphenyl-type compounds conserving features of the fragment molecules were subsequently tested for increased inhibition. Kinetic analysis and molecular modelling provided insight into the putative binding mode of these selective, micromolar inhibitors. Preliminary *in vivo* tests against adherent mouse fibroblast 3T6 cells revealed weak cytotoxicity.

## Methods

*Materials.* Dimethylsulfoxide (DMSO) was purchased from Fisher Scientific (Fair Lawn, NJ). The origin of the screened fragments is indicated in Table S1. Solvents and reagents for synthesis are commercially available (Aldrich, Acros Organics, Fisher Scientific, Sigma Chemical Co.) and were used without further purification. Dihydrofolate (DHF) was synthesized from folic acid as described<sup>22</sup>.  $\beta$ -NADPH was purchased from Alexis biochemicals (San Diego, CA). Unless otherwise indicated, other reagents for enzyme expression and assays were from BioShop Canada inc. (Burlington, ON).

*Synthesis of symmetrical bis-benzamidine, bis-benzimidazoles and bis-p-hydroxybenzoate.* Compounds **8** and **9** were prepared and characterized following a known procedure<sup>23</sup>. Briefly, a mixture of a 4-hydroxybenzaldehyde (5.12g, 42 mmol), a dibromoalkane (C<sub>4</sub> or C<sub>5</sub>) or dibromopropan-2-ol (20 mmol), and potassium carbonate (2.76 g, 20 mmol) in ethanol (10 ml) was heated under reflux for 8h. After cooling, the precipitate was filtered and successively washed with water, ethanol, and ether. The bisbenzaldehyde was sufficiently pure to proceed to the next step. A mixture of the bisbenzaldehyde (3 mmol), sodium pyrosulfite (0.57 g, 3 mmol), 3,4-diaminobenzoic acid (0.91 g, 6 mmol), and water (3 ml) in ethanol (9 ml) was microwave irradiated (Biotage) for 15 min at 140°C. After cooling, the precipitate was filtered and thoroughly washed with water, ethanol, and ether. Yield: **8**, 97 %; **9**, 94 %. The synthesis of compound **8b** was the same except for substitution of the diaminobenzoic acid with *ortho*-phenylenediamine (6 mmol)<sup>23</sup>. The synthesis of compound **8a** has been reported<sup>24</sup>, however a slightly modified version was performed. In short, a mixture of a 4-hydroxybenzoic acid (0.69 g, 5 mmol), dibromopentane (7.5 mmol), potassium hydroxide (0.84 g, 15 mmol), and water (1.5 ml) in ethanol (13.5 ml) was microwave irradiated for 20 min at 120°C. After cooling, the precipitate was filtered and thoroughly washed with water, ethanol, and ether. Yield: 72 %. Compounds were characterized by IR and <sup>1</sup>H-NMR (see Supporting Information).

*Enzyme purification.* Recombinant Type II R67 DHFR<sup>18, 25</sup> with an *N*-terminal 6-histidine tag was overexpressed in *Escherichia coli* BL21 containing plasmid pRep4 (Qiagen). Fresh overnight culture (3 mL) was added to Terrific Broth (1 L; 100 µg/mL ampicillin and 50 µg/mL kanamycin) and the flask was incubated at 37 °C and 250 rpm until OD<sub>600</sub> reached approximately 0.7. Protein expression was induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a concentration of 1 mM, with further incubation for 3 h. The cells were harvested by centrifugation (30 min, 4000g, 4 °C). The cell pellet was resuspended in 30 mL of lysis buffer (0.1 M potassium phosphate, 5 mM imidazole, pH 8.0), and the cells were disrupted by one passage through a cell disrupter (Constant Systems) adjusted to 27 kpsi. An additional 10 mL of buffer washed residual lysate through. Following centrifugation (30 min, 23500 g, 4 °C) and filtration on a 0.22 µm filter, the supernatant was injected onto a 5 mL His-Trap HP cartridge at a flow-rate of ~1 mL/min using an Äkta FPLC (GE Healthcare). The column was washed with ~12 CV of lysis buffer. A linear gradient (6 CV) of the same buffer + imidazole from 0 to 30 mM with a plateau (6 CV) at 30 mM was followed by a step to 300 mM imidazole for elution. Fractions containing R67 DHFR were identified according to activity assay and analysis on tricine-SDS-PAGE<sup>26</sup>, and pooled for dialysis at 4°C into 0.1 M phosphate buffer, pH 8.0 using 3500 Da molecular weight cut-off dialysis tubing (Spectrum Laboratories). Protein concentration was determined using the Bradford assay protein assay (Bio-Rad) using bovine serum albumin (Bio-Rad) as a protein standard. Human chromosomal DHFR was overexpressed in *E. coli* BL21 (DE3) and purified as described previously<sup>27</sup>.

*Fragment-based inhibitor screening against purified enzymes.* Substrates were quantified by spectrophotometry in 50 mM potassium phosphate pH 8.0 ( $\epsilon_{340\text{nm}} = 6\,200\text{ M}^{-1}\text{cm}^{-1}$  for NADPH and  $\epsilon_{282\text{nm}} = 28\,400\text{ M}^{-1}\text{cm}^{-1}$  for DHF). For inhibitor screening, enzyme activity was determined by monitoring the depletion of NADPH and DHF at 340 nm ( $\Delta\epsilon_{340\text{nm}} = 12\,800\text{ M}^{-1}\text{cm}^{-1}$ ) in 400mM HEPES buffer, pH 7.0 for R67 DHFR and 50 mM potassium phosphate pH 8.0 for hDHFR. The high concentration of HEPES was required to offer adequate buffering in the presence of high concentrations of acidic or

basic compounds, because the activity of R67 DHFR is exquisitely sensitive to pH changes<sup>28</sup>.

To provide adequate compound dissolution in a solvent system where both DHFRs retained appreciable activity, the maximal solubility of each test compound was determined in HEPES-buffered 10 % DMSO. In that system, R67 DHFR retained 31 % activity and hDHFR retained 17 % activity, allowing reliable assessment of inhibition. For inhibition assays, each test compound was dissolved in neat DMSO and taken up in the reaction buffer containing the substrates, to a final DMSO concentration of 10 %. For assay automation, testing was performed in a 96-well plate format with a reaction volume of 200  $\mu$ L. When test compounds contributed to an excessive background absorbance at 340 nm, activity was monitored in a volume of 100  $\mu$ L to reduce the path length. Liquid handling was carried out using a BioMek NX automated workstation (Beckman Coulter) and data collected with a Beckman DTX 880 plate-reader. Unless otherwise noted, final concentrations of NADPH and DHF were 50  $\mu$ M. As a result, NADPH is at  $30 \times K_M^{\text{NADPH}}$ <sup>18</sup>. The inclusion of 10% DMSO weakens DHF binding, increasing apparent  $K_M^{\text{DHF}}$  approximately 3-fold<sup>29</sup>, which is taken into account in all calculations. Here, DHF is at  $\sim 2 \times K_M^{\text{DHF}}$ . These conditions provide a clear spectrophotometric signal in the case of uninhibited activity, without masking potential competitive inhibition. Reactions were initiated with addition of 250 nM R67 DHFR. Initial rates (generally less than 10 % substrate conversion) were determined. Determination of selectivity using hDHFR was performed under the same conditions save the buffer (see above). The concentration of hDHFR was 40 nM to allow measurement of initial rates while compensating for the higher turnover constant of hDHFR relative to R67 DHFR.

*Determination of binding and kinetic parameters.* The IC<sub>50</sub>s were determined with GraphPad Prism 5 using the log[inhibitor] vs response equation; K<sub>i</sub> values were determined using the Cheng-Prusoff equation:  $K_i = \text{IC}_{50}/([S]/K_M^{\text{NADPH}} + 1)$ , assuming a competitive mode of inhibition<sup>30</sup>.  $K_M^{\text{NADPH}}$  used for the calculation was  $1.6 \pm 0.02 \mu\text{M}$ <sup>18</sup>. For determination of the mode of inhibition, the method of Dixon was used, in 50 mM potassium phosphate buffer pH 7.0 with 10 % DMSO. For inhibition relative to



NADPH, DHF was held at 50  $\mu$ M, NADPH was at 5  $\mu$ M or 80  $\mu$ M (3-fold and 50-fold  $K_M^{\text{NADPH}}$ , respectively) and inhibitor over a range of concentrations spanning  $K_i$ . For inhibition relative to DHF, NADPH was held at 50  $\mu$ M, DHF was at 25  $\mu$ M or 164  $\mu$ M ( $\sim K_M^{\text{DHF}}$  and  $\sim 7$ -fold  $K_M^{\text{DHF}}$ ) with inhibitor again spanning  $K_i$ . Initial rates were determined in 1 cm path-length cuvettes using a Cary 100 Bio UV–visible spectrophotometer (Agilent). For determination of binding stoichiometry, NADPH was held at 16  $\mu$ M and DHF at 80  $\mu$ M (10-fold  $K_M^{\text{NADPH}}$  and  $\sim 3$ -fold  $K_M^{\text{DHF}}$ , respectively) and titrated with increasing concentrations of **9**. Binding data were analyzed according to the Hill equation for inhibition:  $\log(v_i/(v_0-v_i)) = -n \log[I_i] + \log K_d$ , where  $v_i$  is the velocity at inhibitor concentration  $[I_i]$ ,  $v_0$  is the non-inhibited velocity,  $n$  is the Hill coefficient giving an indication of the number of binding sites and  $K_d$  is the dissociation constant.

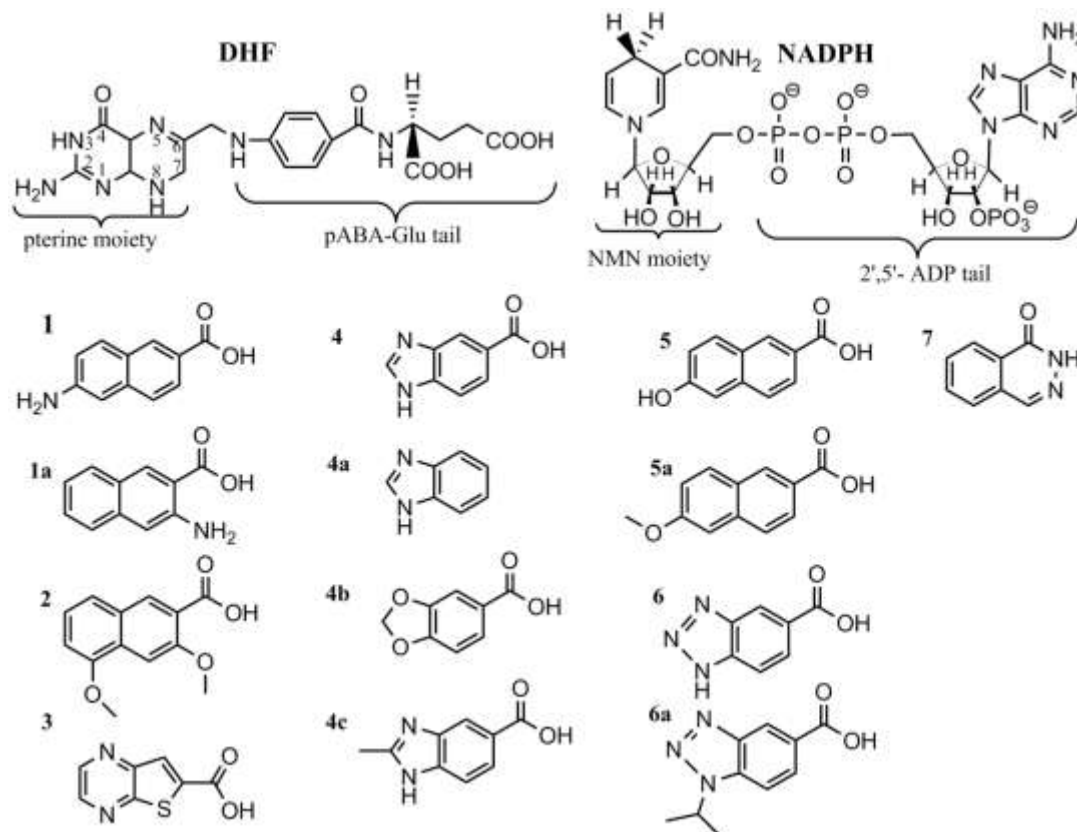
*Molecular docking simulations.* The crystal structure of a ternary complex of R67 DHFR (pdb file 2RK1<sup>19</sup>) was used for all docking simulations. The ligand 3D structures were drawn using Chemdraw 3D Pro 8.0. The geometry of the ligands was optimized using HyperChem 8.0.3 with the default settings. Docking simulations were run on a Dell XPS 1645 with an i7-720QM processor and 6 GB DDR3 running either Molegro Virtual Docker 4.0.0<sup>31</sup> or Autodock Vina 1.1.1<sup>32</sup>. For Molegro, the Moldock Score [Grid] scoring function was used. Grid resolution was 0.30 Å with a radius of 30. The search algorithm used was Moldock Optimizer, with default settings. The number of runs varied from 10 to 25 and the maximum iterations varied from 2000 to 10 000. For Autodock Vina, the default settings were kept except the exhaustiveness parameter which was increased from 8 to 25. Structural visualization was performed with Moldock.

*Eukaryotic cell proliferation assay.* The cytotoxicity of compounds **8**, **8a**, **8b** and **9** was evaluated using 3T6 fibroblast cells (kindly provided by the Laboratory of Biology and Embryology of the University of Mons-Hainaut) in a eukaryotic cell viability test performed with the MTT (3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide) reduction assay<sup>33</sup>. MTT is a water soluble tetrazolium salt that is cleaved into an insoluble purple formazan by the succinate dehydrogenase system

of the mitochondrial chain, which is active only in live cells. 3T6 Fibroblasts are a permanent mouse embryonic fibroblast cell line. Cells were cultured in DMEM supplemented with 10 % heat-inactivated FBS (Invitrogen), in a CO<sub>2</sub> incubator (37°C, 5% CO<sub>2</sub>). Trypsinized 3T6 cells (200 µl,  $5 \times 10^3$  cells) were seeded in each well of a 96-well plate (except for the wells on the edge). Eight successive dilutions of each compound (in DMSO, Sigma Aldrich) were realized, from 200 µM to 0.05 µM. After 24 h of incubation, the culture medium was replaced with 200 µl of fresh medium containing the same compound dilutions. Culture medium (200 µl) was also added to the cell growth control rows while the solvent control wells received 200 µl of the relevant mixture of DMSO and culture medium. Blanks (no cells) received the same mixture. After a further 48 hours of incubation, the cells were carefully washed with PBS and 100 µl of fresh culture medium was added into each well. MTT (100 µl) (*In vitro* toxicology assay kit, MTT based, Sigma Aldrich) was then added to all wells, and incubated for 3 hours. The reaction was stopped by adding MTT solubilization reagent. After overnight incubation at 37°C, the formazan dye was quantified with a microplate reader (Thermo Labsystems Multiskan Ascent 354) using a test wavelength of 540 nm and a reference wavelength of 690 nm. Eight concentrations of each test compound (200 µM to 0.05 µM) were verified in triplicate. The % inhibition of cell proliferation was calculated as follows: % IC =  $100 - [\text{corrected mean OD sample} \times 100 / \text{corrected mean OD solvent controls}]$  where % IC = % inhibition of cell proliferation, and corrected mean OD sample/solvent =  $\text{mean OD}_{540-690} \text{ of samples/controls} - \text{mean OD}_{540-690} \text{ of blanks}$ . For each compound, the % inhibition of activity was plotted against the compound concentration scale. The concentration inhibiting 50 % of cell proliferation (IC<sub>50</sub>) was the x-axis value corresponding to one-half of the maximal absorbance value.

## Results and discussion

### *Fragment-based inhibitor screening of R67 DHFR.*



**Figure 1. Structures of the weak inhibitor fragments and their inactive analogues.** Structures of the DHF substrate and NADPH cofactor are provided as a reference. Compounds: **1**, 6-aminoisonaphthoic acid; **1a**, 3-aminoisonaphthoic acid; **2**, 3,5-dimethoxyisonaphthoic acid; **3**, thieno[2,3-b]pyrazine-6-carboxylic acid; **4**, 1H-benzimidazole-5-carboxylic acid; **4a**, benzimidazole; **4b**, piperonylic acid; **4c**, 2-methyl-1H-benzimidazole-5-carboxylic acid; **5**, 6-hydroxy-2-naphthoic acid; **5a**, 6-methoxy-2-naphthoic acid; **6**, 1H-benzotriazole-5-carboxylic acid; **6a**, 1-methylethyl-1H-benzotriazole-5-carboxylic acid; and **7**, 1(2H)-phthalazinone. No inhibition was observed up to 5 (**1a** and **5a**), 10 (**4c** and **6a**), or 30 mM (**4a** and **4b**), which are the highest concentrations that could be tested due to constraints of poor solubility and high absorbance.

To date, two non-specific micromolar inhibitors of R67 DHFR have been reported: novobiocin ( $K_i = 60 \mu\text{M}$ ), a DNA gyrase inhibitor, and congo red, a diazo dye ( $K_i = 2 \mu\text{M}$ )<sup>17</sup>. Their structural differences relative to the other known micromolar ligands of R67 DHFR, namely the DHF substrate and NADPH co-factor, as well as the loose substrate specificity of the enzyme<sup>15</sup>, immediately suggest that the active-site is amenable to fragment-based inhibitor design. Accordingly, we selected a panel of small

organic compounds (fragments) for inhibitor screening, based on structural similarity to the moieties of the substrate and co-factor which bind deepest in the active-site pore<sup>19</sup>. Approximately 100 commercially-available compounds which share some chemical features of the DHF substrate pteroyl group or NADPH cofactor nicotinamide ring were screened for selective inhibition of R67 DHFR relative to hDHFR (Table S1). Because the compounds are simple and possess a low molecular weight (150 to 250 g/mol), no strong inhibition was expected. Hence, they were tested for inhibition at mM concentrations. Figure 3.1 illustrates compounds **1-7**, which provided inhibition with an IC<sub>50</sub> in the low- to mid-millimolar range (Table 3.1), as well as some structural analogs which provided no inhibition.

**Table 3.1. Inhibition of R67 DHFR and hDHFR by fragments 1-7 <sup>a</sup>.**

Compound	IC <sub>50</sub>			K <sub>i</sub>			Relative activity			LE
	(mM)			(μM)			hDHFR (%) <sup>b</sup>			(kcal)
1	1.7	± 0.3		51	± 9		14	± 13		-0.32
2	1.8	± 0.4		55	± 14		14	± 4		-0.32
3	2.0	± 1.3		61	± 41		37	± 9		-0.48
4	2.6	± 1.4		80	± 42		75	± 5		-0.47
5	6.1	± 2.3		190	± 71		13	± 2		-0.36
6	8.4	± 2.6		260	± 80		69	± 13		-0.41
7	17	± 1.7		530	± 54		97	± 3		-0.32

<sup>a</sup> Values are given as the average +/- standard deviation from the mean.

<sup>b</sup> Compounds were tested at the highest possible concentration (1 = 2.5 mM, 2 and 3 = 5 mM, 4 and 5 = 10 mM, 6 and 7 = 30 mM), taking into account constraints due to poor solubility and high absorbance.

Compounds **1**, **2** and **5** possess a naphthalene core substituted with a carboxylic acid, but are differently functionalized. Despite the differences between the 6-amino group of **1** and the 2,5-methoxy substitution of **2**, their IC<sub>50</sub> are unchanged (1.7 mM and

1.8 mM for compounds **1** and **2**, respectively). However, with the amino group at position 2 (**1a**), no inhibition was observed. The vicinal carboxylate and amino groups of **1a** have the potential for intramolecular H-bonding which would modulate the electronic effect of these functional groups relative to **1**. Compound **5** is hydroxylated at position 6 and shows inhibition similar to **1** and **2** ( $IC_{50} = 6.1$  mM), while a 6-methoxy substituent abrogated inhibition (**5a**). These results suggest the positive contribution of a hydrophilic substituent at C5 or C6 of the naphthoic acid, opposite the carboxylic acid, though different substituents modulate the inhibitory activity of the naphthoic acid framework.

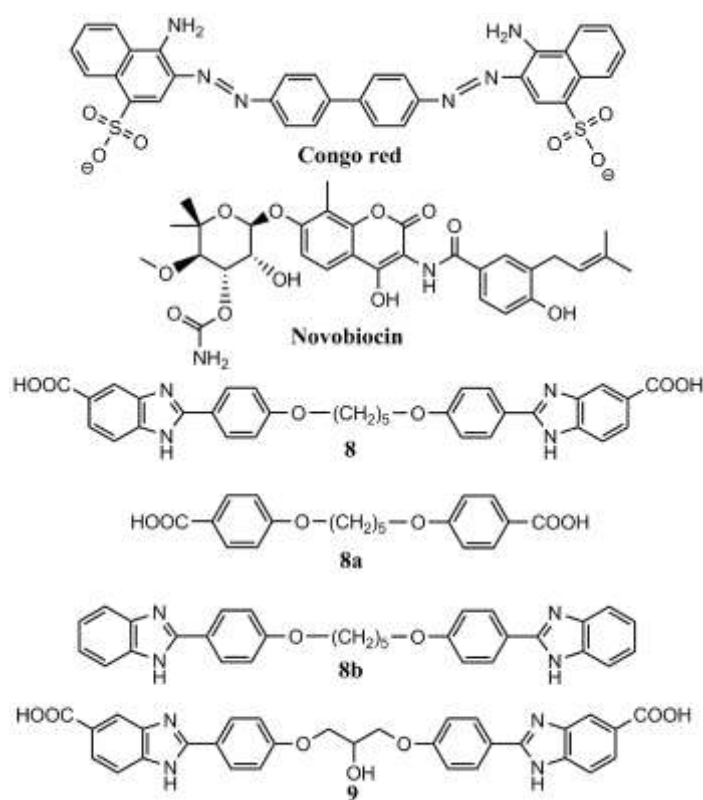
Compound **3**, a thienopyrazine framework substituted with a carboxylic acid, inhibited R67 DHFR with an  $IC_{50}$  of 2.0 mM, in the same range as the similarly substituted benzimidazole **4** ( $IC_{50} = 2.6$  mM) and the carboxylate-substituted benzotriazole **6** ( $IC_{50} = 8.4$  mM). Benzimidazole (**4a**) gave no inhibition, confirming the contribution of the carboxylic acid substitution to inhibition. Changing the aromatic diimino ring to a dioxole ring (compound **4b**) prevented inhibition, suggesting the importance of aromaticity and/or polarity of heteroatoms. It is not clear why an additional methyl substituent at C6 of the acid-substituted benzimidazole (**4c**) abrogated inhibition, since it will be shown below that C6 was used to build fragment **4** into a more complex inhibitor (*vide infra*). An additional isopropyl at N5 of the acid-substituted benzotriazole (**6a**), prevented inhibition, indicating that additional bulk or hydrophobicity at positions opposite the carboxylic acid substituent are to be avoided. The only inhibitor identified that carried no carboxylic acid substituent was **7**, the weakest inhibitor with  $IC_{50} = 17$  mM.

By way of fragment screening, we thus identified several weak inhibitors of R67 DHFR. All are hydrophobic, aromatic and possess a carboxylic acid (except the weaker inhibitor **7**). The positions opposite the carboxylic acid modulate the activity of the compound and could be targeted to improve these inhibitors.

To direct inhibitor development toward selectivity for R67 DHFR while sparing human DHFR (hDHFR), we tested these **7** fragments against hDHFR. Compounds **4** and **6** displayed good selectivity: at  $3 \times IC_{50}$ , approximately 80% of hDHFR activity

remained. Compound **7** did not affect the activity of the hDHFR up to 30 mM ( $2 \times \text{IC}_{50}$ ). However, compounds **1**, **2**, **3**, **5** inhibited non-selectively: at  $< 2 \times \text{IC}_{50}$ , these compounds completely inhibited hDHFR. Though **4** and **6** were not the most potent, they showed a good selectivity for R67 DHFR and they have a good ‘ligand efficiency’ (LE), where LE is  $\Delta G$  of binding (calculated from  $K_i$ ) divided by the number of heavy atoms in the molecule (Table 3.1). LE indicates how efficiently a compound binds to its target, relative to its molecular weight. In a recent example, careful optimization of millimolar fragments with a LE of -0.30 kcal and less, provided nanomolar inhibitors<sup>21</sup>. Fragment **4** possesses a good potency, selectivity and a LE of -0.47 kcal; it was thus selected as a basis for development of a more potent inhibitor.

*Inhibition of R67 DHFR by symmetrical polyaromatic molecules.*



**Figure 3.2. Structures of the R67 DHFR inhibitors and their inactive analogs.** Congo red and novobiocin are non-specific inhibitors of R67 DHFR. **8**: 2,2'-[1,5-pentanediy]bis(4-oxyphenylene)]-1H-benzimidazole-5-carboxylic acid. **8a**: 4,4'-[1,5-pentanediy]-(oxy)]-bisbenzoic acid. **8b**: 2,2'-[1,5-pentanediy]bis(4-oxyphenylene)]-1H-benzimidazole. **9**: 2'-(4,4'-(2-hydroxypropane-1,3-diyl)-bis-(oxy)-bis-(4,1-phenylene))-bis-1H-benzimidazole-5-carboxylic acid.

As mentioned above, novobiocin ( $K_i = 70 \mu\text{M}$ ) and congo red ( $K_i = 2 \mu\text{M}$ ) (Figure 3.2) are the only 2 known R67 DHFR inhibitors and neither is specific to R67 DHFR. Both structures are elongated and possess substituted aromatic rings. Notably, the more potent congo red is symmetrical, with a negatively charged substituent on each extremity and a central hydrophobic linker. R67 DHFR possesses a symmetrical and mostly hydrophobic active site<sup>25</sup>, the center of which holds the reactive groups of the DHF substrate and NADPH cofactor. Upon binding of DHF and NADPH, Lys32 from the four DHFR protomers – which lie at the mouth of the active-site channel – bind to the negatively-charged DHF carboxylates and NADPH phosphates<sup>34</sup>. We hypothesized that the sulfites of congo red and carboxylic acid substituents of fragments **1** to **6** interact with these lysines, or with the neighbouring Lys33 (from each of the four protomers), which lie further outside of the active-site channel. Thus, we tested inhibition with symmetrical compounds constituted of two fragments **4** symmetrically connected with a semi-rigid hydrophobic linker (yielding **8**) or a shorter, slightly more hydrophobic linker (yielding **9**) (Figure 3.2).

**Table 3.2. Inhibition of R67 DHFR and hDHFR by symmetrical compounds **8** and **9** <sup>a</sup>.**

Compound	IC <sub>50</sub>			K <sub>i</sub> <sup>b</sup>			Relative activity			LE
	(μM)			(μM)			hDHFR (%) <sup>c</sup>			(kcal)
<b>8</b>	64	±	11	2.0	±	0.3	95	±	11	-0.18
<b>8a</b>	≥ 1000			N.D			100	±	1	N.D.
<b>8b</b>	≥ 300			N.D			110	±	14	N.D.
<b>9</b>	130	±	11	4.0	±	0.3	95	±	10	-0.18

<sup>a</sup> Values are given as the average +/- standard deviation from the mean.

<sup>b</sup> K<sub>i</sub> was calculated from IC<sub>50</sub> values.

<sup>c</sup> Compounds were tested at the highest possible concentration (**8** = 400 μM, **8a** = 1 mM, **8b** = 300 μM, **9** = 1 mM), taking into account constraints due to poor solubility and high absorbance. N.D.: Not determined

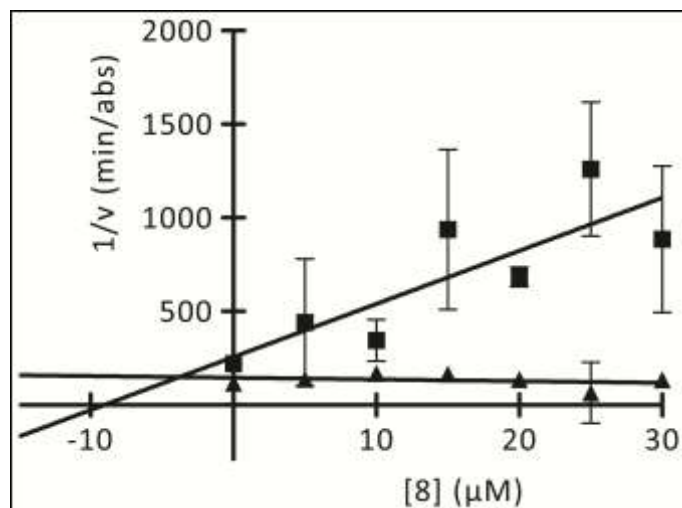
We observed an improvement in potency, with  $IC_{50}$  of 64  $\mu$ M and 130  $\mu$ M for compounds **8** and **9** respectively (compared to 2.6 mM for **4**) (Table 3.2 and Figure S1). To gain insight into the structure-activity relation (SAR), we tested two analogs of **8** (**8a** and **8b**). Removing the benzimidazole moiety from **8** (yielding **8a**) provided no inhibition up to the highest concentration tested (1 mM). Removing the carboxylic acid from the benzimidazole moiety (yielding **8b**) reduced the potency: at 300  $\mu$ M, more than 60% of R67 activity remained, demonstrating that the carboxylic acid functions are important contributors to binding. Nonetheless, lack of inhibition with **8a** demonstrated that carboxylic acids alone are not sufficient for binding to R67 DHFR. To verify if the selectivity of the inhibitors was conserved, we tested inhibitors **8** and **9** against hDHFR (Table 3.2). No inhibition of hDHFR was observed up to the highest concentrations tested (400  $\mu$ M of **8**, 1 mM of **8a**, 300  $\mu$ M of **8b** and 1 mM of **9**).

Overall, starting from fragment **4**, which gave a millimolar-range  $IC_{50}$  and a LE of -0.47 kcal (Table 3.1), we identified two selective inhibitors of R67 DHFR (**8** and **9**) which showed  $IC_{50}$  300-600 fold lower, in the micromolar range. Compounds **8** and **9** possess a LE of -0.18 kcal (Table 3.2), and thus bind less efficiently relative to their molecular weight than the precursor fragment **4**. This illustrates that they are poorly optimized compounds. Thus, inhibitors **8** and **9** provide a starting point to develop potent and selective inhibitors against R67 DHFR.

#### *Inhibition is competitive.*

To understand the binding mode of the symmetrical inhibitors to their target, the mode of inhibition of **8** toward R67 DHFR was determined according to the method of Dixon, by performing inhibition at varying NADPH and inhibitor concentrations<sup>35</sup>. Figure 3.3 illustrates that **8** is a competitive inhibitor, with  $K_i = 4.2 \mu$ M against the cofactor NADPH. This value is consistent with that calculated from  $IC_{50}$  (Table 3.2). Compound **8** showed similar results toward the substrate DHF, with a  $K_i$  of 9.7  $\mu$ M (Figure S1). The competitive nature of the inhibition with respect to both the substrate and the cofactor is consistent with the inhibitor binding within the active site cavity.





**Figure 3.3.** Dixon plot for determination of type of inhibition of R67 DHFR by **8**, relative to the NADPH cofactor. The reciprocal rates of DHFR activity were plotted as a function of inhibitor concentration. DHF substrate was held at 50  $\mu\text{M}$ . NADPH cofactor was held at 5  $\mu\text{M}$  ( $3 \times K_M^{\text{NADPH}}$ , ■) or 80  $\mu\text{M}$  ( $50 \times K_M^{\text{NADPH}}$ , ▲). Values are given as the mean  $\pm$  standard deviation for triplicate results. The intercept of the two slopes gave  $K_i(8) = 4.2 \mu\text{M}$ .

#### *Docked model of inhibitor binding to R67 DHFR.*

To help direct future efforts in improving the inhibitors, compounds **8**, **8a**, **8b** and **9** were docked into the crystal structure of R67 DHFR (2RK1<sup>19</sup>). Each of the four protomers of the R67 DHFR homotetramer plays a different role when a molecule binds in the active site, thus breaking the tetramer's symmetry<sup>36</sup>. For clarity, we will refer to residues according to their chain identity (protomer<sup>a</sup> to protomer<sup>d</sup>), as required. Figure 3.4 (A-C) shows that R67 DHFR forms a symmetrical homotetramer where the active site lies within the central channel. The entrances to the pore are lined with the positively-charged Lys32 and Lys33, each contributed by symmetry-related protomers (*i.e.* protomer<sup>a</sup> and protomer<sup>d</sup>, or protomer<sup>b</sup> and protomer<sup>c</sup>) to provide four lysines at each entrance to the pore (Figure 3.4C). They serve to bind and orient the phosphates of the cofactor NADPH and the carboxylates of the substrate DHF<sup>34</sup>. The NADPH nicotinamide ring and the DHF pteroyl group, which contain the hydride donor and acceptor, respectively, meet at the mostly hydrophobic centre of the pore for reactivity,

where Gln67, Ile68 and Tyr69 are key binding residues. In particular, Gln67 forms a H-bond network with Tyr69, providing a molecular ‘clamp’ that holds the substrates in place<sup>19</sup>. The unusual symmetry of the active site pore and its binding promiscuity<sup>15</sup>, in addition to the cooperativity of NADPH and DHF binding<sup>37</sup>, increase the difficulty of obtaining meaningful docking results. Docking to the apo-form of the crystallized ternary complex was shown to give results that were inconsistent with that complex<sup>38</sup>, though docking of one ligand in the presence of the second has proven more reliable<sup>38-39</sup>. Nonetheless, we assume that inhibitor binding follows the simplest model of binding to the apoenzyme – this is consistent with inhibition being competitive against both DHF and NADPH, though it does not necessarily demonstrate it. DHF binding was not deemed a meaningful control because only its pterin ring has been resolved. We thus conducted a docking control on the apoenzyme using the ligand NADPH. In the five poses corresponding to the best docking scores, the nicotinamide moiety and the three phosphates were roughly well localized, although the adenine ring made contacts with different faces of the channel. Thus the docking of NADPH to the apoenzyme could not be directly compared to the bound conformer in the crystallized ternary complex but served as a guide for relative docking scores of the inhibitors.

Compounds **8**, **8a**, **8b** and **9** were docked onto R67 DHFR using Molegro Virtual Docker. The docking scores are reported in Table 3.3, with NADPH as the reference compound.  $K_M^{NADPH}$  is 1.6  $\mu M$ <sup>18</sup> (for the ternary complex). Its docking score (for the binary complex) was -190, where the score is a dimensionless unit estimating the binding energy. The scores for inhibitors **8** and **9** were similar, at -190 and -185, respectively, consistent with their  $K_i$  (for **8**: calculated from  $IC_{50} = 2.0 \mu M$ , experimentally determined = 4.2  $\mu M$ ; for **9**: calculated from  $IC_{50} = 4.0 \mu M$ ) which are similar to  $K_M^{NADPH}$ . Compounds **8a** and **8b** provided weaker scores of -109 and -156, respectively, reflecting the poor inhibition properties that were apparent in their respective  $IC_{50}$ s.

**Table 3.3. Docking of NADPH and 8, 8a, 8b and 9 to R67 DHFR.**

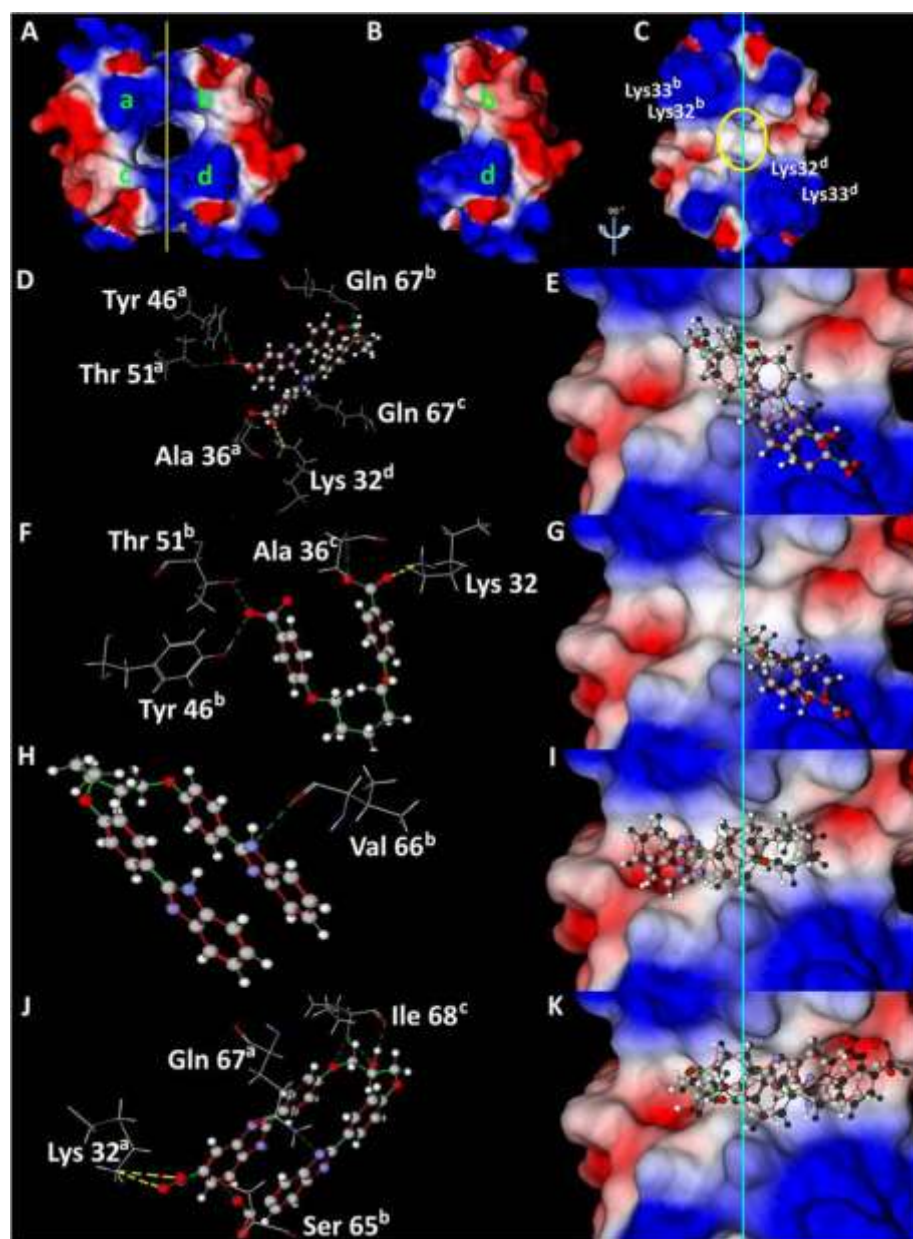
Compound	Moldock Score
NADPH	-190
<b>8</b>	-190
<b>8a</b>	-109
<b>8b</b>	-156
<b>9</b>	-185

Typical poses obtained for these compounds are shown in Figure 3.4 (D-K). Interestingly, each of the four R67-bound inhibitors was predicted to adopt a U-shape, with stacking of the symmetry-related aromatic moieties. The aromatic stacking would fill the volume of the cavity in a manner similar to the juxtaposition of the NADPH and DHF, during the native reaction. This proposed binding mode is consistent with the symmetry of the tetrameric active site, where each arm of the inhibitor forms similar contacts with symmetry-related residues of the enzyme (*i.e.* protomer<sup>a</sup> and protomer<sup>d</sup>). Further, it is consistent with the hydrophobic nature of the center of the active site, which would disfavour ‘threading’ of the charged extremity of any of these compounds through the active site channel. Overall, it is consistent with a competitive mode of inhibition. Specific features of the predicted mode of binding are discussed below.

Compound **9** has the most hydrophilic central spacer; interestingly, the docked model predicts that **9** is deeply bound within the active site, where the glycerol spacer appears to strongly interact with the backbone of Ile68<sup>c</sup> (Figure 3.4J-K). Ile68 is proposed to provide recognition of the carboxamide on the nicotinamide ring, and the amide on the pteridine ring<sup>19</sup>. As a result, the benzene and benzimidazole would form Van der Waals interactions with the central hydrophobic pocket. The central pocket would be fully occupied, as shown in Figure 3.4 (K) where the docked molecule is predicted to traverse the center of symmetry of the enzyme. Lys32<sup>a</sup> has the potential to form an electrostatic bond with a benzimidazole carboxylate (Figure 3.4J), consistent with its role in binding the NADPH phosphate or the DHF glutamate tail to orient these

molecules in the active site. Moreover, in the model Gln67<sup>a</sup> H-bonds with the benzimidazole moiety and Ser65<sup>b</sup> H-bonds with the terminal carboxylate.

The docked model of compound **8** (Figure 3.4D-E) was bound less deeply within the hydrophobic pocket, although it was also predicted to traverse the lengthwise center of symmetry of the channel. Interactions with the backbone of Ile68 were lost due to the more hydrophobic nature of the spacer, but interactions with Gln67<sup>b</sup> and Gln67<sup>c</sup> were gained. This position allows the terminal carboxylates to H-bond with residues Thr51<sup>a</sup>, Tyr46<sup>a</sup> and Ala36<sup>a</sup>, in addition to allowing electrostatic interaction with Lys32<sup>d</sup>. Compound **8b** (Figure 3.4H-I) differs from **8** by its lack of terminal carboxylates on the benzimidazole moieties. **8b** was also deeply docked in the active site. The interactions with Thr51, Tyr46, Ala36 and Lys32 predicted with **9** were lost with **8b** due to the lack of the carboxylates.



**Figure 3.4: Docking of 8, 8a, 8b and 9 R67 DHFR using Molegro Virtual Docker.** (A) R67 DHFR is shown as an electrostatics surface with the basic, acidic and hydrophobic regions shown in blue, red and white respectively, and protomers labelled a-d. (B) Two protomers are shown, where the tetramer was split as indicated in panel A. (C) A 90° counter-clockwise rotation of panel B, revealing the largely hydrophobic active site cavity (circled). (D to K) Docking of 8, 8a, 8b and 9. (D-E) Docking of 8; (F-G) 8a; (H-I) 8b; (J-K) 9. The left-hand panels show putative contacts formed, with the compound in ball-and-stick representation and the residues of R67 DHFR in sticks representation. Hydrogen, carbon, nitrogen and oxygen are in white, grey, blue and red balls, respectively. Red and green bonds represent the non-rotatable and rotatable bonds, respectively, as defined for docking experiments. Yellow and green dashed lines represent electrostatic bonds and H-bonds, respectively. The right-hand panels show the same result, with the enzyme active site oriented as in panel C. The vertical blue line indicates the center of symmetry for the axis that runs lengthwise through the binding channel.

This result is consistent with the  $IC_{50}$  for **8b**, which is a least 5-fold higher than **8**. Compound **8a** (Figure 3.4F-G) differs from **8-9** by its lack of benzimidazole moieties. In the docked model, the terminal carboxylates formed the same H-bonding and electrostatic interactions observed with **8**, which suggests that the benzimidazole moiety is not essential to the formation of these interactions. Putative interactions with Gln67 and Ile68 were lost as a consequence of the loss of the benzimidazole heteroatoms and the replacement of the glycerol moiety by pentane, which could explain the lower docking score and the lack of *in vitro* inhibition. Compound **8a** was also the least deeply docked into the active site.

Docking with Autodock Vina afforded essentially superimposable results (Figure S3): the inhibitors folded into a U-shape, slipping the centre of the molecule deep into the active site channel with the carboxylates generally held by Lys32; Tyr46, Thr51 and Ser65 also participated in carboxylate binding (not shown). Similar interactions were predicted with Gln67, Ile68 and Tyr69 as when docking with Molegro.

While these models suggest that intramolecular stacking occurs upon binding, this may result from limitations of the docking algorithms used. Specifically, docking algorithms dock a single molecule of test compound at a time. Thus, the U-shape we observed may reflect a virtual solution to maximize molecular packing within the active-site tunnel. An alternative docking method of compounds **8**, **8a** and **9** has shown a linear mode of binding (H. Hogue, E. Purisima and T. Sulea, pers. comm.). In that case, the molecule threads through the length of the active site pore, forming essentially the same contacts as in the U-shape, except that the two R67 DHFR protomers it interacts with are on opposite ends of the channel (thus interacting with protomer<sup>a</sup> and protomer<sup>c</sup>, rather than protomer<sup>a</sup> and protomer<sup>d</sup>). Indeed, it is possible that two molecules of a given compound bind simultaneously, laying flat on top of each other, thus maximizing the contacts with the large active site tunnel. That mode of binding remains consistent with the hydrophobicity of the tunnel, with the putative role of the four-fold Lys32 residues in binding the inhibitors' carboxylates, with the volume occupied by the stacking of the natural DHF substrate and NADPH cofactor and provides a more entropically-favoured solution to binding; threading of the terminal carboxylate through

the hydrophobic active site may be promoted by the water molecules known to populate the active-site pore<sup>19</sup>. Attempts were made to soak inhibitor **9** into R67 DHFR crystals that were obtained using a recently described crystallization protocol<sup>25</sup>. A change in the electron density was observed in the active site pore in the presence of **9** relative to the apo structure of R67 DHFR (B. Yachnin, D.Colin, A. Berghuis, pers. comm.). Unfortunately, because the pore lies on the crystallographic symmetry axes, combined with the fact that the ligand did not appear to be present at full occupancy, it proved impossible to interpret the electron density. From what could be seen, the density was not consistent with the U-shaped conformation of compound **9** as suggested by the dockings results, but the poor quality of the electron density precludes drawing conclusions on the actual bound conformation of the inhibitor.

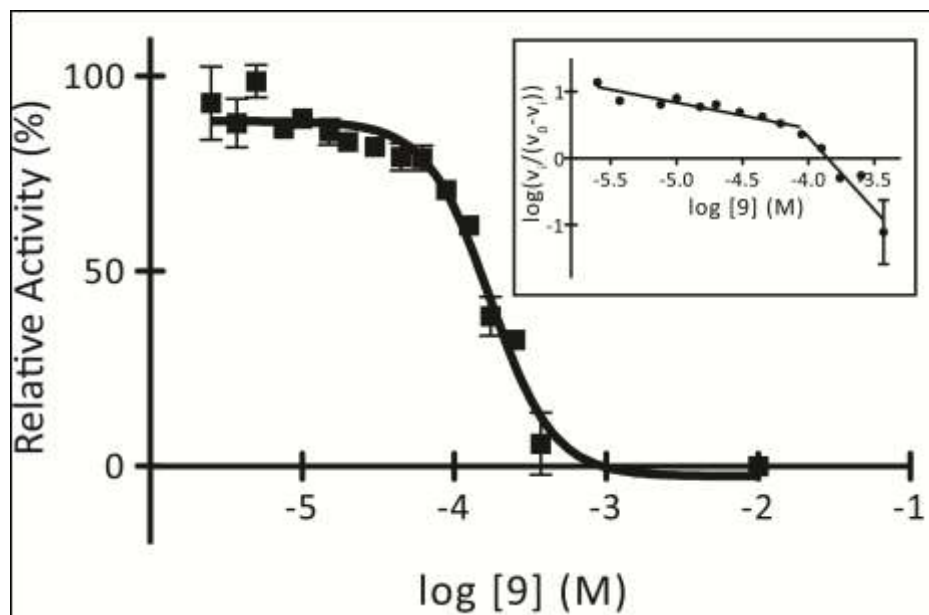
Notwithstanding the precise mode of binding, symmetry appear to play a key role in binding and selectivity for R67 DHFR. The proposed interaction of terminal carboxylates with Lys32, Thr51, Tyr46 or Ala36 may be modulated by the distance between the carboxylate and the hydrophobic pocket as observed by comparing docked models of **8** and **9**. Interaction with the backbone of Ile68 and Val66 may be promoted by a heteroatom near the hydrophobic pocket.

#### *Binding stoichiometry.*

As mentioned above, if the inhibitor molecules **8** or **9** were to bind to the active site in a linear conformation, it is conceptually possible for two molecules to bind simultaneously. In contrast, the model of the U-shaped conformation of bound **8** or **9** would not allow this (Figure 3.4). We examined binding stoichiometry by performing competitive inhibition under conditions appropriate for determining a Hill coefficient. As shown in Figure 3.5 (inset), the Hill plot with **9** shows a clear inflection, consistent with more than one inhibitor molecule binding to the target<sup>40</sup>. The apparent Hill coefficients obtained are  $0.39 \pm 0.06$  at low inhibitor concentrations and  $2.2 \pm 0.4$  at high inhibitor concentration, while the Hill coefficient obtained upon plotting in semi-logarithmic format is  $2.1 \pm 0.3$  (Figure 3.5). These values are consistent with R67

having the potential to bind more than one inhibitor molecule simultaneously<sup>40</sup>. However, there was no evidence of binding cooperativity according to a replot of reaction rate relative to inhibitor concentration (not shown)<sup>40</sup>.

These results support binding of the inhibitors in a linear mode, where it is sterically possible to stack two molecules into the active site, rather than the U-shaped docked models (Figure 3.4) where inhibitors **8** and **9** both cross the center of symmetry of the active site such that it would not be possible to fit in two molecules simultaneously. It should be noted that a further sterically plausible binding mode would be for one molecule of inhibitor to bind in a linear mode concurrently with one molecule of NADPH or DHF. We attempted to verify this according to fluorescence binding measurements, but compounds **8** and **9** both gave rise to strong fluorescence plagued by intense quenching at the relevant micromolar concentrations.



**Figure 3.5: Hill plot for determination of the stoichiometry of binding of compound 9 to R67 DHFR.** Semi-logarithmic representation of the relative activity vs the inhibitor concentration. DHF substrate was held at 80  $\mu\text{M}$  ( $\sim 3 \times K_M^{\text{DHF}}$ ) and NADPH cofactor at 16  $\mu\text{M}$  ( $10 \times K_M^{\text{NADPH}}$ ). The calculated Hill coefficient was  $2.1 \pm 0.3$ . The inset represents the corresponding double logarithmic Hill plot. The best fit to the data does not follow a single straight line but is a composite of two straight lines with different slopes. The apparent Hill coefficients were  $0.39 \pm 0.06$  for low inhibitor concentrations and  $2.2 \pm 0.4$  for high inhibitor concentrations. Values are given as the mean  $\pm$  standard deviation for triplicate results.



*In vivo assays.*

Determination of the minimum inhibitory concentration (MIC) for compounds **8** and **9** was attempted using *E. coli* XL1-Blue expressing DHFR R67 in minimum media (M9) containing 50 µg/mL of TMP (to inhibit the native, chromosomal *E. coli* DHFR). Bacterial growth was not inhibited up to 1mM of **8** or **9** (maximal concentration tested; data not shown), suggesting that these compounds do not penetrate into *E. coli*. We considered testing esterified versions of the inhibitors to increase their permeability in microbial cells, but their solubility became too low to allow testing at meaningful concentrations. Nonetheless, **8b** has an antiproliferative effect on *Leishmania donovani*, *Pneumocystis carinii*, *Plasmodium falciparum* and *Trypanosoma brucei rhodesiense*, though their mode of action has not yet been determined<sup>23</sup>. Compounds **8**, **8a**, **8b** and **9** exhibited only weak cytotoxicity against adherent mouse fibroblast 3T6 cells with the exception of compound **8b**, which exhibited an IC<sub>50</sub> of approximately 1 µM (Table 3.4).

**Table 3.4. Cytotoxicity of 8, 8a, 8b and 9 toward 3T6 fibroblast cells.**

Compound	IC <sub>50</sub> (µM)
<b>8</b>	40 ± 6% mortality at 200 µM
<b>8a</b>	34.0 ± 0.4
<b>8b</b>	0.98 ± 0.07
<b>9</b>	40 ± 5% mortality at 200 µM

## Conclusions

We have developed a functional screening platform for identification of selective inhibitors of the bacterial antibiotic-resistance enzyme, R67 DHFR. The platform is based on semi-automated determination of enzymatic activity in presence of a variety of fragments, or simple compounds similar to the native ligands of the enzyme, to assess inhibition. We observed that a variety of small aromatic compounds offer mM-range inhibition of R67 DHFR, which is consistent with the proposed ‘primitive’ nature of its relatively promiscuous binding site<sup>15,39</sup>. Weakly inhibiting molecules provided the basis for testing compounds of greater complexity, which provided an increase in affinity from millimolar to micromolar. Importantly, we screened the human DHFR in parallel with R67 DHFR to identify structures with the best prospects for selective inhibition. Thus, we identified a new class of selective, symmetrical and competitive inhibitors of R67 DHFR (**8** and **9**), which are as potent as congo red but are selective toward R67 DHFR. Compounds **8** and **9** were both based on 2-benzimidazolecarboxylic acid (**4**), which had an IC<sub>50</sub> of 3 mM (Table 1). Joining two 2-benzimidazolecarboxylic acid units into symmetrical structures with a central, hydrophobic and flexible linker, increased affinity to the low micromolar range. Testing of the analogs **8a** and **8b** demonstrated that carboxylic acid groups and benzimidazoles contribute to affinity. Preliminary cytotoxicity and selectivity experiments demonstrated that **8** and **9** are not highly toxic, thus can serve as a lead compounds to develop more potent selective inhibitors of R67 DHFR. We are attempting to gain structural information to determine the mode of binding of the bis-benzimidazolium compounds to R67 DHFR, to direct further efforts in discovery and development of selective inhibitors for this emerging drug-resistance target. Future inhibitors should take advantage of the important structural differences between R67 DHFR and human DHFR. In particular, the large size and symmetry of the R67 DHFR active site appears to allow simultaneous binding of two inhibitor molecules in a manner that is not coherent with the narrower, asymmetric active site cleft of human DHFR<sup>41</sup>. This work provides inspiration for the design of the next generation of inhibitors.

## Acknowledgments

The authors thank the Combinatorial Chemistry Laboratory at l'Université de Montréal and the Medicinal Chemistry Platform at the Institut de Recherche en Immunologie et en Cancérologie (IRIC) de l'Université de Montréal for access to compounds. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), as well as PROTEO, the Québec Network for Research on Protein Structure, Function and Engineering, and CGCC, the Québec Centre for Green Chemistry and Catalysis, which are funded by les Fonds Québécois pour la Recherche sur la Nature et les Technologies (FQRNT).

**Supporting Information Available:** The full list of fragments screened against R67 DHFR, characterization of compounds **8**, **8a** and **9** as well as additional kinetic and molecular docking results are provided. This material is available free of charge *via* the Internet at <http://pubs.acs.org>.

## References

1. Huovinen, P.; Sundstrom, L.; Swedberg, G.; Sköld, O., Trimethoprim and sulfonamide resistance. *Antimicrob Agents Chemother* **1995**, 39 (2), 279-289.
2. (a) Huovinen, P., Resistance to trimethoprim-sulfamethoxazole. *Clin Infect Dis* **2001**, 32 (11), 1608-1614; (b) Forgacs, P.; Wengenack, N. L.; Hall, L.; Zimmerman, S. K.; Silverman, M. L.; Roberts, G. D., Tuberculosis and trimethoprim-sulfamethoxazole. *Antimicrob Agents Chemother* **2009**, 53 (11), 4789-4793.
3. 

[http://www.who.int/selection\\_medicines/committees/expert/18/applications/Sulfamethoxazole\\_trimethoprim.pdf](http://www.who.int/selection_medicines/committees/expert/18/applications/Sulfamethoxazole_trimethoprim.pdf) (accessed November 29, 2011).
4. Kehrenberg, C.; Schwarz, S., Trimethoprim resistance in a porcine *Pasteurella aerogenes* isolate is based on a *dfrA1* gene cassette located in a partially truncated class 2 integron. *J Antimicrob Chemother* **2011**, 66 (2), 450-452.
5. Cabello, F. C., Heavy use of prophylactic antibiotics in aquaculture: a growing problem for human and animal health and for the environment. *Environ Microbiol* **2006**, 8 (7), 1137-1144.
6. Le, T. X.; Muneke, Y.; Kato, S., Antibiotic resistance in bacteria from shrimp farming in mangrove areas. *Science of The Total Environment* **2005**, 349 (1-3), 95-105.
7. Segura, P. A.; Garcia-Ac, A.; Lajeunesse, A.; Ghosh, D.; Gagnon, C.; Sauve, S., Determination of six anti-infectives in wastewater using tandem solid-phase extraction and liquid chromatography-tandem mass spectrometry. *Journal of Environmental Monitoring* **2007**, 9 (4), 307-313.
8. (a) Burrige, L.; Weis, J. S.; Cabello, F.; Pizarro, J.; Bostick, K., Chemical use in salmon aquaculture: A review of current practices and possible environmental effects. *Aquaculture* **2010**, 306 (1-4), 7-23; (b) Medina, A.; Horcajo, P.; Jurado, S.; De La Fuente, R.; Ruiz-Santa-Quiteria, J. A.; Domínguez-Bernal, G.; Orden, J. A., Phenotypic

and Genotypic Characterization of Antimicrobial Resistance in Enterohemorrhagic Escherichia Coli and Atypical Enteropathogenic E. Coli Strains from Ruminants. *Journal of Veterinary Diagnostic Investigation* **2011**, 23 (1), 91-95; (c) Barlow, R. S.; Pemberton, J. M.; Desmarchelier, P. M.; Gobius, K. S., Isolation and characterization of integron-containing bacteria without antibiotic selection. *Antimicrobial Agents and Chemotherapy* **2004**, 48 (3), 838-842; (d) Zhao, H. X.; Shen, J. Z.; An, X. P.; Fan, H. L.; Cao, J. S.; Li, P. F., Characterization of integrons in multiple antimicrobial resistant Escherichia coli isolates from bovine endometritis. *Res Vet Sci* **2011**, 91 (3), 412-414.

9. Volpato, J. P.; Pelletier, J. N., Mutational 'hot-spots' in mammalian, bacterial and protozoal dihydrofolate reductases associated with antifolate resistance: sequence and structural comparison. *Drug Resist Updat* **2009**, 12 (1-2), 28-41.

10. (a) Plowe, C. V.; Kublin, J. G.; Dzinjalama, F. K.; Kamwendo, D. S.; Mukadam, R. A.; Chimpeni, P.; Molyneux, M. E.; Taylor, T. E., Sustained clinical efficacy of sulfadoxine-pyrimethamine for uncomplicated falciparum malaria in Malawi after 10 years as first line treatment: five year prospective study. *BMJ* **2004**, 328 (7439), 545; (b) Musset, L.; Pradines, B.; Parzy, D.; Durand, R.; Bigot, P.; Le Bras, J., Apparent absence of atovaquone/proguanil resistance in 477 Plasmodium falciparum isolates from untreated French travellers. *J Antimicrob Chemother* **2006**, 57 (1), 110-115.

11. Appleman, J. R.; Prendergast, N.; Delcamp, T. J.; Freisheim, J. H.; Blakley, R. L., Kinetics of the formation and isomerization of methotrexate complexes of recombinant human dihydrofolate reductase. *J Biol Chem* **1988**, 263 (21), 10304-10313.

12. Sköld, O., Resistance to trimethoprim and sulfonamides. *Vet Res* **2001**, 32 (3-4), 261-273.

13. (a) Fling, M. E.; Walton, L.; Elwell, L. P., Monitoring of plasmid-encoded, trimethoprim-resistant dihydrofolate reductase genes: detection of a new resistant enzyme. *Antimicrob Agents Chemother* **1982**, 22 (5), 882-888; (b) Stone, D.; Smith, S. L., The amino acid sequence of the trimethoprim-resistant dihydrofolate reductase specified in Escherichia coli by R-plasmid R67. *J Biol Chem* **1979**, 254 (21), 10857-10861.

14. Alonso, H.; Gready, J. E., Integron-sequestered dihydrofolate reductase: a recently redeployed enzyme. *Trends Microbiol* **2006**, *14* (5), 236-242.
15. Holland, J. C.; Linn, C. E.; DiGiammarino, E.; Nichols, R.; Howell, E. E., Does R67 dihydrofolate reductase possess a proton donor? *Adv Exp Med Biol* **1993**, *338*, 493-498.
16. Huovinen, P., Trimethoprim resistance. *Antimicrob Agents Chemother* **1987**, *31* (10), 1451-1456.
17. Howell, E. E., Searching sequence space: two different approaches to dihydrofolate reductase catalysis. *Chembiochem* **2005**, *6* (4), 590-600.
18. Schmitzer, A. R.; Lépine, F.; Pelletier, J. N., Combinatorial exploration of the catalytic site of a drug-resistant dihydrofolate reductase: creating alternative functional configurations. *Protein Eng Des Sel* **2004**, *17* (11), 809-819.
19. Krahn, J. M.; Jackson, M. R.; DeRose, E. F.; Howell, E. E.; London, R. E., Crystal structure of a type II dihydrofolate reductase catalytic ternary complex. *Biochemistry* **2007**, *46* (51), 14878-14888.
20. (a) Zartler, E. R.; Shapiro, M. J., Fragonomics: fragment-based drug discovery [Review]. *Current Opinion in Chemical Biology* **2005**, *9* (4), 366-370; (b) de Kloe, G. E.; Bailey, D.; Leurs, R.; de Esch, I. J., Transforming fragments into candidates: small becomes big in medicinal chemistry. *Drug Discov Today* **2009**, *14* (13-14), 630-646; (c) Gozalbes, R.; Carbajo, R. J.; Pineda-Lucena, A., Contributions of computational chemistry and biophysical techniques to fragment-based drug discovery. *Curr Med Chem* **2010**, *17* (17), 1769-1794.
21. Murray, C. W.; Rees, D. C., The rise of fragment-based drug discovery. *Nat Chem* **2009**, *1* (3), 187-192.
22. Blakley, R. L., Crystalline Dihydropteroylglutamic Acid. *Nature* **1960**, *188*, 231 - 232.

23. Mayence, A.; Pietka, A.; Collins, M. S.; Cushion, M. T.; Tekwani, B. L.; Huang, T. L.; Eynde, J. J. V., Novel bisbenzimidazoles with antileishmanial effectiveness. *Bioorganic & Medicinal Chemistry Letters* **2008**, 18 (8), 2658-2661.
24. Donahoe, H. B.; Benjamin, L. E.; Fennoy, L. V.; Greiff, D., Synthesis of Potential Rickettsiostatic Agents. 1a I. 4,4'-Dicarboxy- $\alpha,\iota$ -diphenoxyalkanes 1b. *The Journal of Organic Chemistry* **1961**, 26 (2), 474-476.
25. Yachnin, B. J.; Colin, D. Y.; Volpato, J. P.; Ebert, M.; Pelletier, J. N.; Berghuis, A. M., Novel crystallization conditions for tandem variant R67 DHFR yield a wild-type crystal structure. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **2011**; Vol. 67, p 1316-1322.
26. Schagger, H.; von Jagow, G., Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **1987**, 166 (2), 368-379.
27. Volpato, J. P.; Yachnin, B. J.; Blanchet, J.; Guerrero, V.; Poulin, L.; Fossati, E.; Berghuis, A. M.; Pelletier, J. N., Multiple conformers in active site of human dihydrofolate reductase F31R/Q35E double mutant suggest structural basis for methotrexate resistance. *J Biol Chem* **2009**, 284 (30), 20079-20089.
28. Park, H. Y.; Zhuang, P.; Nichols, R.; Howell, E. E., Mechanistic studies of R67 dihydrofolate reductase - Effects of pH and an H62C mutation. *J Biol Chem* **1997**, 272 (4), 2252-2258.
29. Chopra, S.; Dooling, R. M.; Horner, C. G.; Howell, E. E., A Balancing Act between Net Uptake of Water during Dihydrofolate Binding and Net Release of Water upon NADPH Binding in R67 Dihydrofolate Reductase. *J. Biol. Chem.* **2008**, 283 (8), 4690-4698.
30. Cheng, Y.; Prusoff, W. H., Relationship between the inhibition constant ( $K_i$ ) and the concentration of inhibitor which causes 50 per cent inhibition ( $I_{50}$ ) of an enzymatic reaction. *Biochem Pharmacol* **1973**, 22 (23), 3099-3108.

31. Thomsen, R.; Christensen, M. H., MolDock: a new technique for high-accuracy molecular docking. *J Med Chem* **2006**, *49* (11), 3315-3321.
32. Trott, O.; Olson, A. J., AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry* **2010**, *31* (2), 455-461.
33. Mosmann, T., Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J Immunol Methods* **1983**, *65* (1-2), 55-63.
34. Hicks, S. N.; Smiley, R. D.; Hamilton, J. B.; Howell, E. E., Role of ionic interactions in ligand binding and catalysis of R67 dihydrofolate reductase. *Biochemistry* **2003**, *42* (36), 10569-10578.
35. Kakkar, T.; Boxenbaum, H.; Mayersohn, M., Estimation of  $K_i$  in a competitive enzyme-inhibition model: comparisons among three methods of data analysis. *Drug Metab Dispos* **1999**, *27* (6), 756-762.
36. Smiley, R. D.; Stinnett, L. G.; Saxton, A. M.; Howell, E. E., Breaking symmetry: mutations engineered into R67 dihydrofolate reductase, a D2 symmetric homotetramer possessing a single active site pore. *Biochemistry* **2002**, *41* (52), 15664-15675.
37. Bradrick, T. D.; Beechem, J. M.; Howell, E. E., Unusual binding stoichiometries and cooperativity are observed during binary and ternary complex formation in the single active pore of R67 dihydrofolate reductase, a D2 symmetric protein. *Biochemistry* **1996**, *35* (35), 11414-11424.
38. Alonso, H.; Gillies, M. B.; Cummins, P. L.; Bliznyuk, A. A.; Gready, J. E., Multiple ligand-binding modes in bacterial R67 dihydrofolate reductase. *J Comput Aided Mol Des* **2005**, *19* (3), 165-187.
39. Howell, E. E.; Shukla, U.; Hicks, S. N.; Smiley, R. D.; Kuhn, L. A.; Zavodszky, M. I., One site fits both: a model for the ternary complex of folate + NADPH in R67



dihydrofolate reductase, a D2 symmetric enzyme. *J Comput Aided Mol Des* **2001**, *15* (11), 1035-1052.

40. Segel, I. H., *In Enzyme Kinetics : Behavior and Analysis of rapid Equilibrium and Steady-State Enzyme Systems*. Wiley Classics Library ed.; John Wiley and sons: 1993; p 957.

41. Oefner, C.; D'Arcy, A.; Winkler, F. K., Crystal structure of human dihydrofolate reductase complexed with folate. *Eur J Biochem* **1988**, *174* (2), 377-385.

## Supporting Information for:

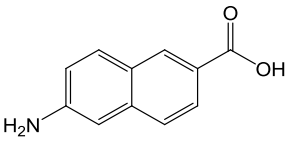
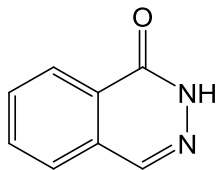
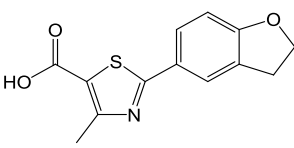
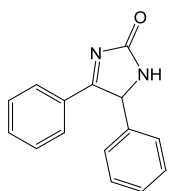
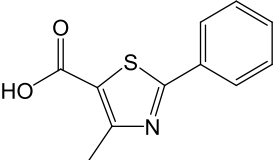
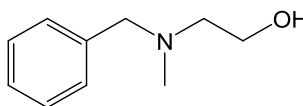
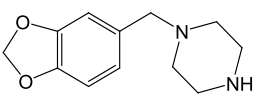
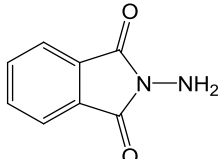
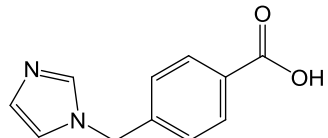
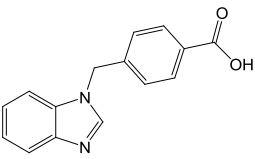
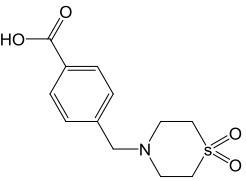
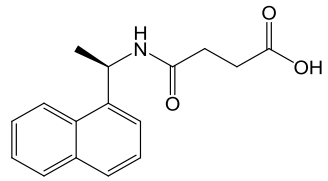
Fragment-based design of symmetrical bis-benzimidazoles as selective inhibitors of the trimethoprim-resistant, Type II R67 dihydrofolate reductase

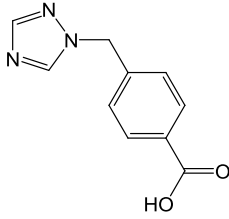
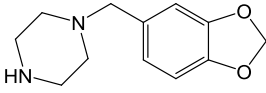
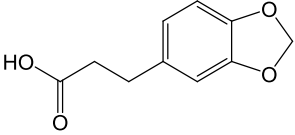
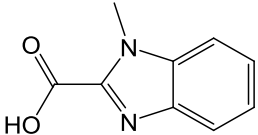
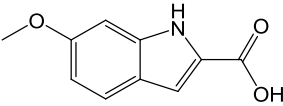
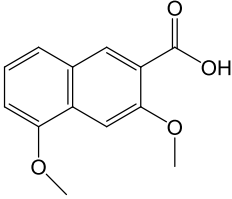
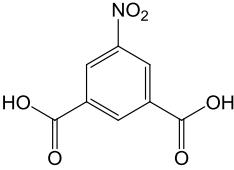
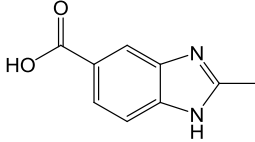
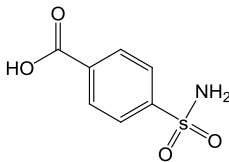
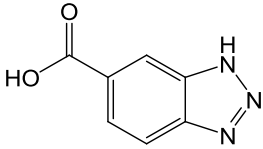
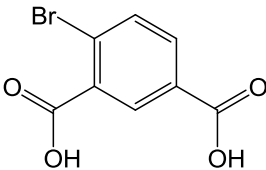
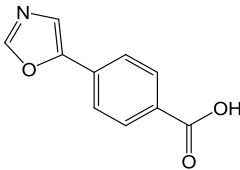
Dominic Bastien, Maximilian C. C. J. C. Ebert, Delphine Forge, Jacynthe Toulouse, Natalia Kadnikova, Florent Perron, Annie Mayence, Tien L. Huang, Jean Jacques Vanden Eynde and Joelle N. Pelletier

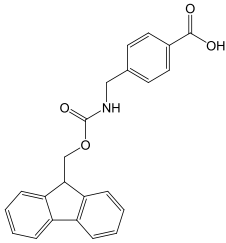
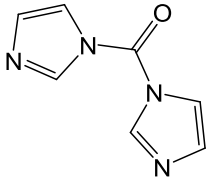
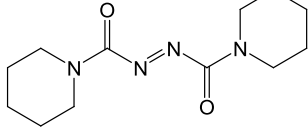
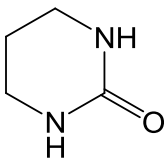
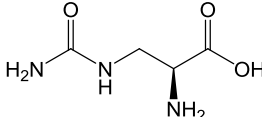
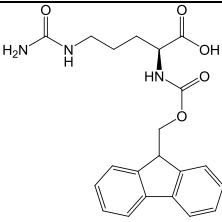
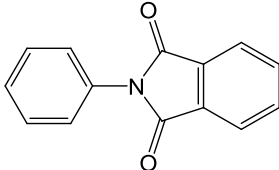
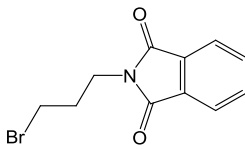
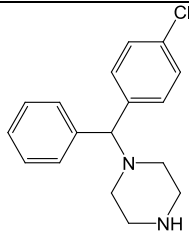
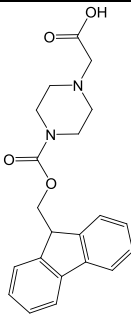
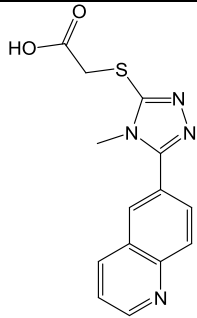
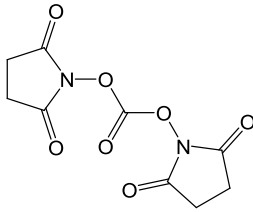
### Table of contents

123	<b>Table S1:</b> Fragments screened against R67 DHFR.
133	Characterization of compounds <b>8</b> , <b>8a</b> and <b>9</b> .
134	<b>Figure S1:</b> IC <sub>50</sub> plot of <b>8</b> and <b>9</b> .
135	<b>Figure S2:</b> Dixon plot for determination of type of inhibition of R67 DHFR by <b>8</b> , relative to the DHF substrate.
136	<b>Figure S3:</b> Docking of compound <b>9</b> on R67 DHFR using AutodockVina.

**Table S1 : List of the compounds screened against R67 DHFR**

#	1	2	3
Structure			
CAS	116668-47-4	119-39-1	690632-04-3
Company	Aldrich	Aldrich	Maybridge
#	4	5	6
Structure			
CAS	642-36-4	33763-20-1	101-98-4
Company	Alfa Aesar	Maybridge	Aldrich
#	7	8	9
Structure			
CAS	32231-06-4	1875-48-5	135611-32-4
Company	Aldrich	Aldrich	Maybridge
#	10	11	12
Structure			
CAS	139742-50-0	465514-21-0	78681-09-1
Company	Maybridge	Not available	Aldrich

#	13	14	15
Structure			
CAS	160388-54-5	32231-06-4	2815-95-4
Company	Maybridge	Aldrich	Maybridge
#	16	17	18
Structure			
CAS	20572-01-4	4382-54-1	98410-68-5
Company	Maybridge	Aldrich	Aldrich
#	19	20	21
Structure			
CAS	618-88-2	709-19-3	138-41-0
Company	Aldrich	Maybridge	Maybridge
#	22	23	24
Structure			
CAS	23814-12-2	6939-93-1	250161-45-6
Company	Maybridge	Not available	Maybridge

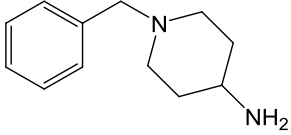
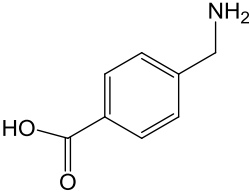
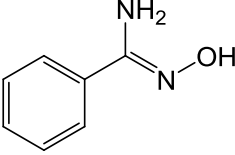
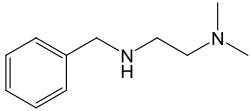
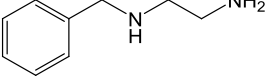
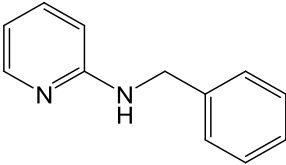
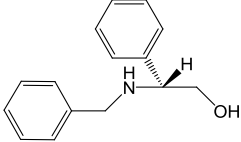
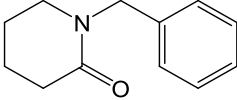
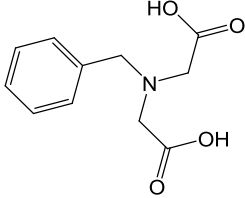
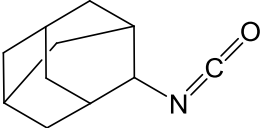
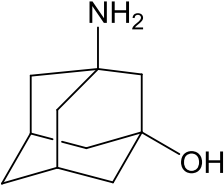
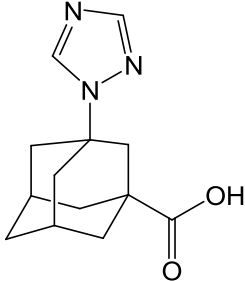
#	25	26	27
Structure			
CAS	164470-64-8	530-62-1	10465-81-3
Company	Chem-Impex	Aldrich	Aldrich
#	28	29	30
Structure			
CAS	1852-17-1	1483-07-4	133174-15-9
Company	Aldrich	Chem-Impex	Adv Chemtech
#	31	32	33
Structure			
CAS	520-03-6	5460-29-7	303-26-4
Company	Aldrich	Aldrich	Aldrich
#	34	35	36
Structure			
CAS	180576-05-0	306935-49-9	74124-79-1
Company	Chem-Impex	Maybridge	Aldrich

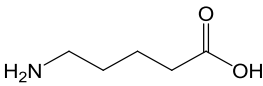
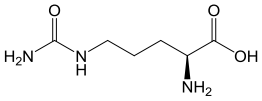
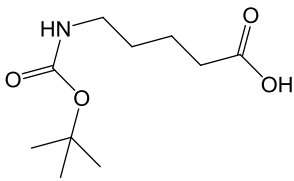
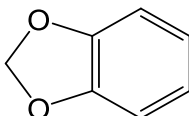
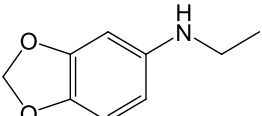
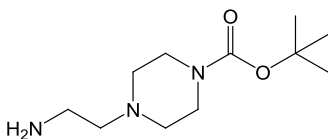
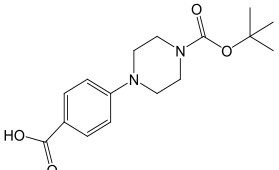
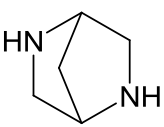
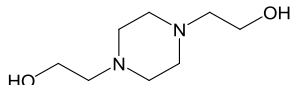
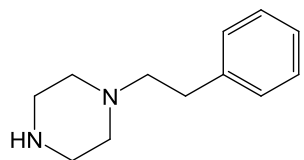
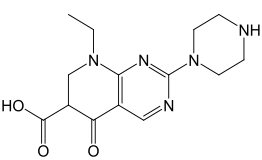
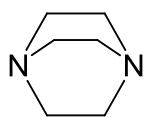
#	37	38	39
Structure			
CAS	2508-29-4	768-94-5	36476-87-6
Company	Aldrich	Aldrich	Maybridge
#	40	41	42
Structure			
CAS	690632-04-3	21256-18-8	118420-05-6
Company	Maybridge	Maybridge	Maybridge
#	43	44	45
Structure			
CAS	76-22-2	34850-66-3	5617-70-9
Company	Aldrich	Aldrich	Aldrich
#	46	47	48
Structure			
CAS	830-13-7	59944-79-5	13139-17-8
Company	Aldrich	Maybridge	Aldrich

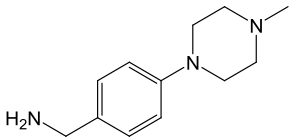
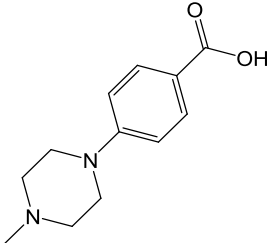
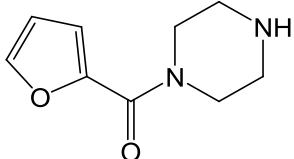
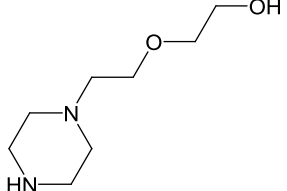
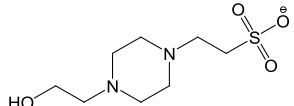
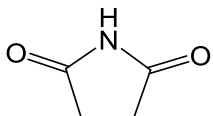
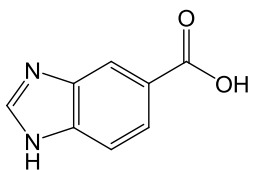
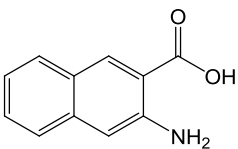
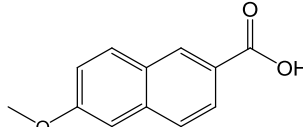
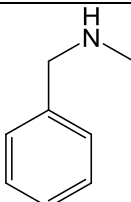
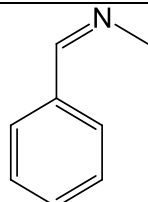
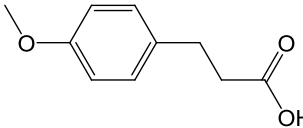
#	49	50	51
Structure			
CAS	104777-39-1	258506-49-9	7113-10-2
Company	Maybridge	Not available	Maybridge
#	52	53	54
Structure			
CAS	22509-74-6	5469-45-4	1485-70-7
Company	Lancaster	Aldrich	Aldrich
#	55	56	57
Structure			
CAS	78818-15-2	500-98-1	16091-26-2
Company	Maybridge	TCI	Trans World
#	58	59	60
Structure			
CAS	Not available	3987-53-9	16154-69-1
Company	Not available	Aldrich	Bionet

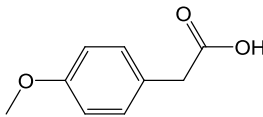
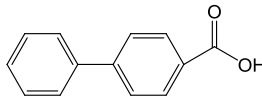
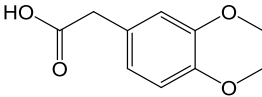
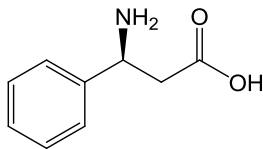
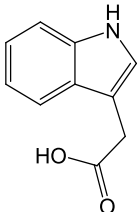
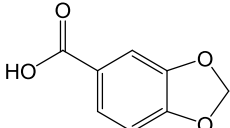
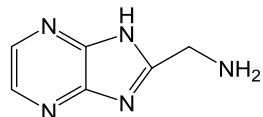




#	73	74	75
Structure			
CAS	50541-93-0	56-91-7	613-92-3
Company	Aldrich	Aldrich	Maybridge
#	76	77	78
Structure			
CAS	103-55-9	4152-09-4	6935-27-9
Company	Lancaster	TCI	Lancaster
#	79	80	81
Structure			
CAS	14231-57-3	4783-65-7	3987-53-9
Company	Aldrich	Aldrich	Aldrich
#	82	83	84
Structure			
CAS	4411-25-0	702-82-9	Not available
Company	Aldrich	Aldrich	Matrix

#	85	86	87
Structure			
CAS	660-88-8	372-75-8	27219-07-4
Company	Aldrich	Aldrich	Aldrich
#	88	89	90
Structure			
CAS	274-09-9	32953-14-3	192130-34-0
Company	Aldrich	TCI	Chembasics
#	91	92	93
Structure			
CAS	162046-66-4	132747-20-7	122-96-3
Company	Jw-PharmLab	Alfa	Aldrich
#	94	95	96
Structure			
CAS	5321-49-3	51940-44-4	280-57-9
Company	Not available	Aldrich	Aldrich

#	97	98	99
Structure			
CAS	216144-45-5	86620-62-4	40172-95-0
Company	Maybridge	Maybridge	Acros
#	100	101	102
Structure			
CAS	13349-82-1	82207-62-3	123-56-8
Company	Aldrich	Fluka	Eastman
#	103	104	105
Structure			
CAS	15788-16-6	5959-52-4	2471-70-7
Company	Aldrich	Aldrich	Aldrich
#	109	110	111
Structure			
CAS	103-67-3	622-29-7	1929-29-2
Company	Aldrich	Aldrich	Aldrich

#	112	113	114
Structure			
CAS	104-01-8	92-92-2	93-40-3
Company	Aldrich	Accurate Chemicals & Scientific	Aldrich
#	115	116	117
Structure			
CAS	40856-44-8	87-51-4	94-53-1
Company	CSPS	Not available	Aldrich
#	118		
Structure			
CAS	5993-91-9		
Company	Aldrich		

**Characterization of newly reported compounds:**

2,2'-[1,5-pentanediy]bis(4-oxyphenylene)]-1H-benzimidazole-5-carboxylic acid; (**8**)

IR(KBr, cm<sup>-1</sup>): 3691-2362, 1680, 1609, 1505, 1256.

<sup>1</sup>H NMR (DMSO-d<sub>6</sub>, ppm): 1.6 (m, 2H, J = 6Hz, CH<sub>2</sub>-CH<sub>2</sub>-CH<sub>2</sub>-O), 1.8 (m, 4H, J = 6Hz, CH<sub>2</sub>-CH<sub>2</sub>-O), 4.1 (t, 4H, J = 6Hz, CH<sub>2</sub>-CH<sub>2</sub>-CH<sub>2</sub>-O), 7.2 (d, 4H, J = 8Hz, C<sub>6</sub>H<sub>4</sub>:H5' and H3'), 7.6 (d, 2H, J = 8Hz, benz: H7), 7.8 (d, 2H, J = 8Hz, benz: H6), 8.2 (m, 6H, C<sub>6</sub>H<sub>4</sub>:H6', H2' and benz: H4), 13 (s, 2H, COOH).

Yield: 97%

4,4'-[1,5-pentanediy]-(oxy)]-bisbenzoic acid(**8a**)

IR(KBr, cm<sup>-1</sup>): 1683, 1605, 1254, 1170, 771.

<sup>1</sup>H NMR (DMSO-d<sub>6</sub>, ppm): 1.5 (m, 2H, J = 6Hz, O-CH<sub>2</sub>-CH<sub>2</sub>-CH<sub>2</sub>), 1.8 (m, 4H, J = 6Hz, O-CH<sub>2</sub>-CH<sub>2</sub>-CH<sub>2</sub>), 4 (t, 4H, J = 6Hz, O-CH<sub>2</sub>-CH<sub>2</sub>-CH<sub>2</sub>), 7 (d, 4H, J = 8 Hz, C<sub>6</sub>H<sub>4</sub>: H3, H5), 7.9 (d, 4H, J = 8Hz, C<sub>6</sub>H<sub>4</sub>: H2, H6), 12.6 (s, 2H, COOH).

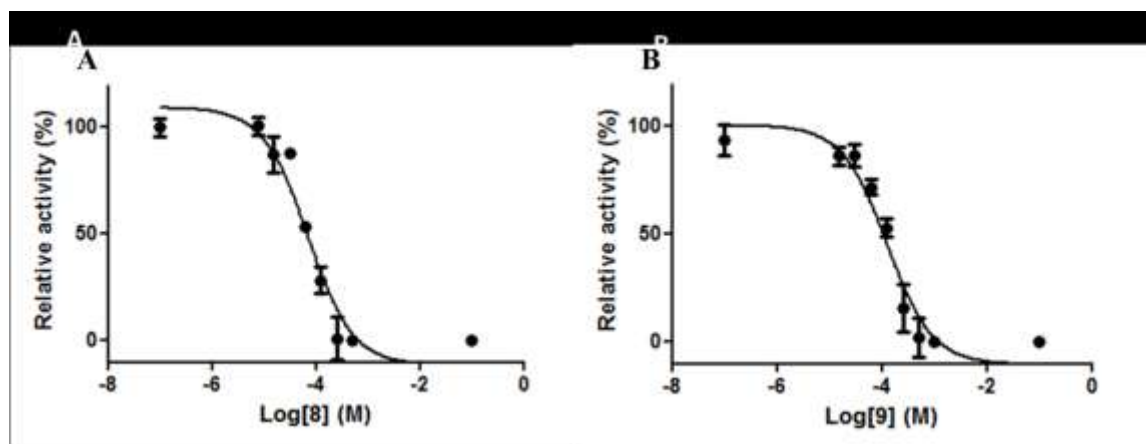
Yield: 72%

2'-(4,4'-(2-hydroxypropane-1,3-diyl)bis(oxy)bis(4,1-phenylene))bis(1H-benzimidazole-5-carboxylic)acid (**9**)

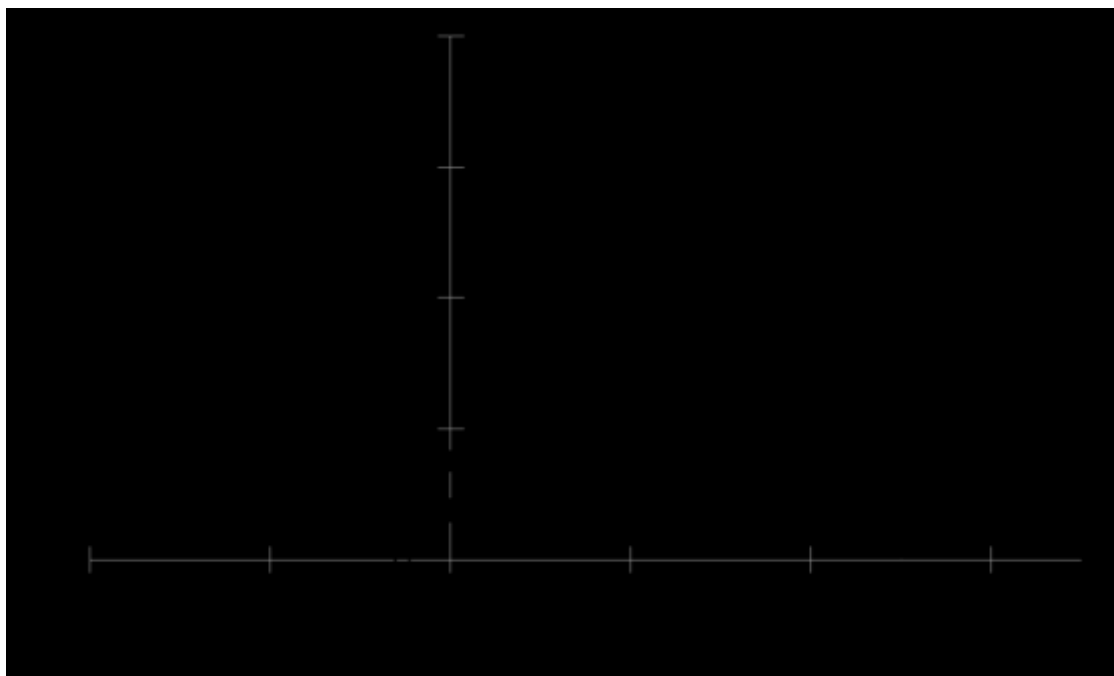
IR(KBr, cm<sup>-1</sup>): 3400, 2950, 1685, 1603.

<sup>1</sup>H NMR (DMSO-d<sub>6</sub>, ppm): 4.24 (m, J = 6Hz, 5H, -CH<sub>2</sub>-CH(OH)-CH<sub>2</sub>), 5.51 (sl, 1H, -OH), 7.22 (d, 4H, J = 8.7 Hz, C<sub>6</sub>H<sub>4</sub>: H3', H5'), 7.66 (d, 2H, benz: H6), 7.86 (d, 2H, benz: H7), 8.18 (m, 6H, C<sub>6</sub>H<sub>4</sub>: H2', H6' and benz: H4), 13.02 (s, 2H, COOH).

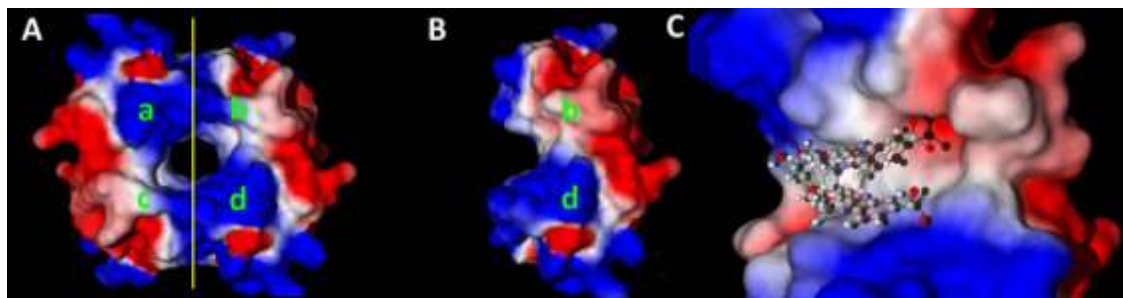
Yield: 94%



**Figure S1: IC<sub>50</sub> plot of 8 and 9.** The activity of R67 DHFR was measured in presence of **8** (A) or **9** (B), and was corrected according to the control without inhibitor. The data were fit to the ‘one-site binding’ equation. Values are given as the mean  $\pm$  standard deviation for triplicate results.



**Figure S2. Dixon plot for determination of type of inhibition of R67 DHFR by 8, relative to the DHF substrate.** The reciprocal rates of DHFR activity were plotted as a function of inhibitor concentration. NADPH co-factor (DHF) was held constant at 50  $\mu\text{M}$ . DHF substrate was held at either 25  $\mu\text{M}$  ( $3 \times K_M$ , ■) or 164  $\mu\text{M}$  ( $15 \times K_M$ , ●).  $K_i = 9.7 \mu\text{M}$ . Values are given as the mean  $\pm$  standard deviation for triplicate results.



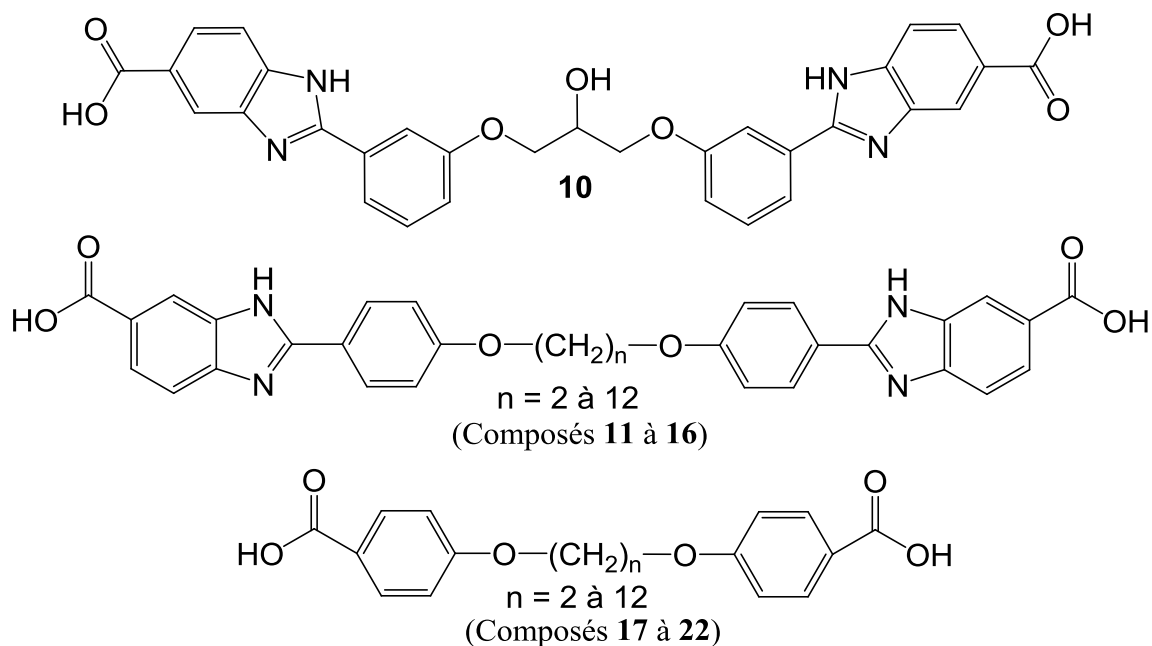
**Figure S3: Docking of compound 9 on R67 DHFR using AutodockVina.** (A) R67 DHFR is shown as an electrostatics surface with the basic, acidic and hydrophobic regions shown in blue, red and white respectively, and monomers labelled a-d. (B) Two monomers are shown, where the tetramer was split as indicated in panel A. (C) Conformation of **9** in the active site of the DHFR R67 shown in ball and sticks representation. Hydrogen, carbon, nitrogen and oxygen are in white, grey, blue and red, respectively.



## Chapitre 4 - Relation structure activité

Au chapitre 3, on a vu que par design à base de fragment, on a été en mesure d'identifier une série de composés pouvant inhiber la DHFR R67 dans l'ordre des bas micromolaires (composés **8** et **9**). De plus, on a démontré l'importance des fonctions carboxyliques (composé **8b**) et du fragment benzimidazole dans ces molécules (composé **8a**).

Afin d'identifier d'autres aspects importants des composés, on a testé plusieurs analogues sur la DHFR R67 (figure 4.1). Le composé **10** permet de vérifier l'importance de l'orientation de l'espaceur sur le groupement phenyle, par rapport à l'orientation (para) des composés **8** et **9**. Les composés **11** à **16** permettent de vérifier l'importance de la longueur de l'espaceur. Finalement, les composés **17** à **22** permettent de vérifier l'importance du groupement benzimidazole et de vérifier si son retrait peut être comblé par un allongement de la molécule.

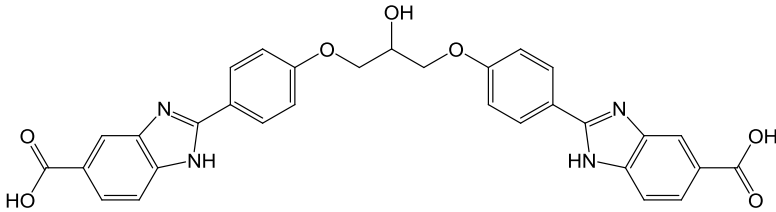
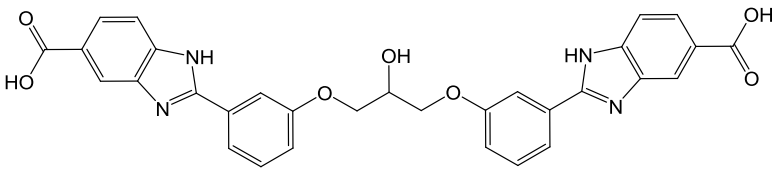


**Figure 4.1:** Analogues des inhibiteurs **8** et **9** qui ont été testés sur la DHFR R67. Les composés **11**, **12**, **13**, **14**, **15** et **16** possèdent un espaceur de 2, 4, 6, 8, 10 et 12 méthylènes respectivement. Les composés **17**, **18**, **19**, **20**, **21** et **22** possèdent un espaceur de 2, 4, 6, 8, 10 et 12 méthylènes respectivement.

## Section 4.1 - Orientation des substituants du phényle

Les inhibiteurs **8** et **9**, identifiés au chapitre 3 possèdent tous deux leur phényle en orientation para par rapport au benzimidazole et à l'espaceur, ce qui donne des molécules linéaires. Deux modèles de liaisons de ces inhibiteurs sur la DHFR R67 ont été abordés au chapitre 3. On propose que les inhibiteurs adoptent une conformation linéaire ou qu'ils se replient en une forme de U dans le site actif de la DHFR R67. Afin de vérifier l'importance de la linéarité des inhibiteurs, le composé **10** fut testé. Ce composé est un analogue de l'inhibiteur **9**, mais son phényle porte le benzimidazole et l'espaceur en position méta au lieu de para ce qui diminue la linéarité de la molécule.

**Tableau 4.1 : Comparaison du IC<sub>50</sub> des composés 9 et 10.**

Numéro composé	Structure	IC <sub>50</sub> (μM)
9		131 ± 11
10		511 ± 73

Le IC<sub>50</sub> du composé **10** est environ 4 fois plus grand que celui du composé **9** (tableau 4.1). Ces résultats démontrent que la position para et/ou que des molécules ayant une géométrie linéaire favorisent la liaison avec la DHFR R67. Bien qu'il ne le démontre pas, ce résultat appuie l'hypothèse que les inhibiteurs adoptent une conformation linéaire dans le site actif de la DHFR puisqu'une orientation méta ne devrait pas affecter grandement la capacité de la molécule de se replier en U. Par contre, l'orientation méta réduit la distance entre les deux carboxylates proposés pour établir des interactions électrostatiques avec les Lys32 situées à l'embouchure du site actif.

## Section 4.2 - Effet de la longueur de l'espaceur

Dans le but de vérifier l'importance de la longueur des inhibiteurs, les composés **11** à **16** furent testés. La figure 4.2 illustre l'effet de la longueur de ces composés sur le  $IC_{50}$ . On observe un creux énergétique quand l'espaceur possède de 4 à 8 méthylènes. Lorsqu'on sort de cet intervalle, on constate que le  $IC_{50}$  augmente rapidement. Ainsi, pour ce type de molécule, un espaceur de 4 à 8 méthylènes est idéal.

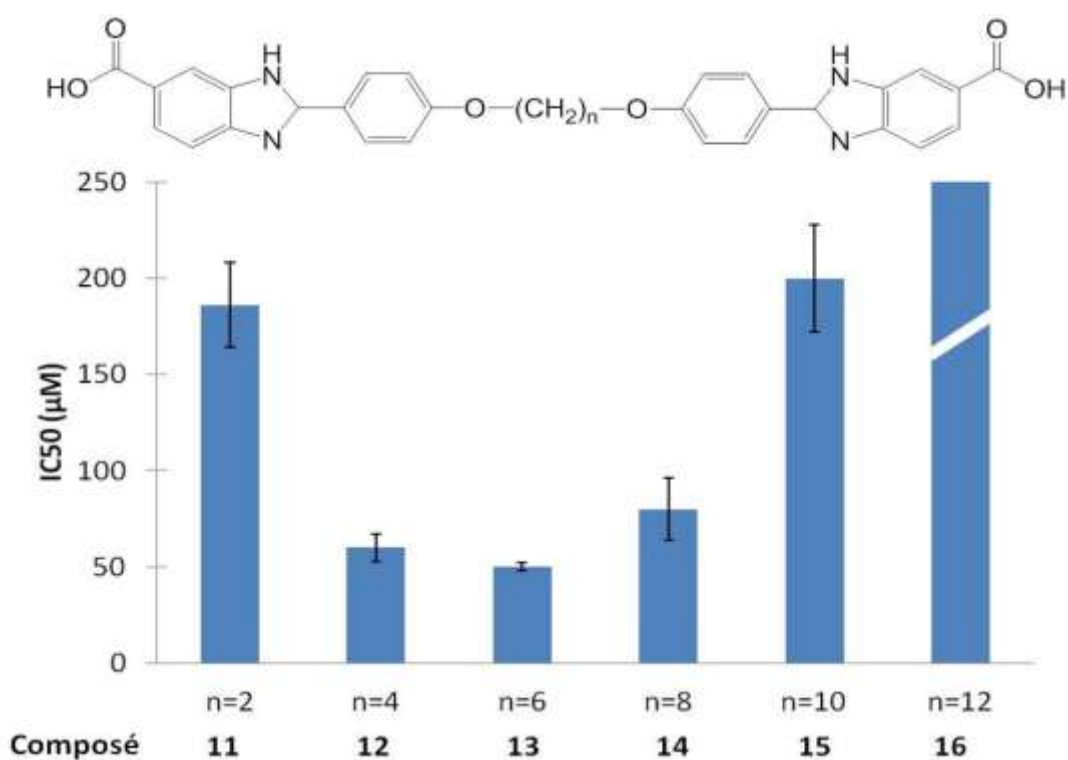
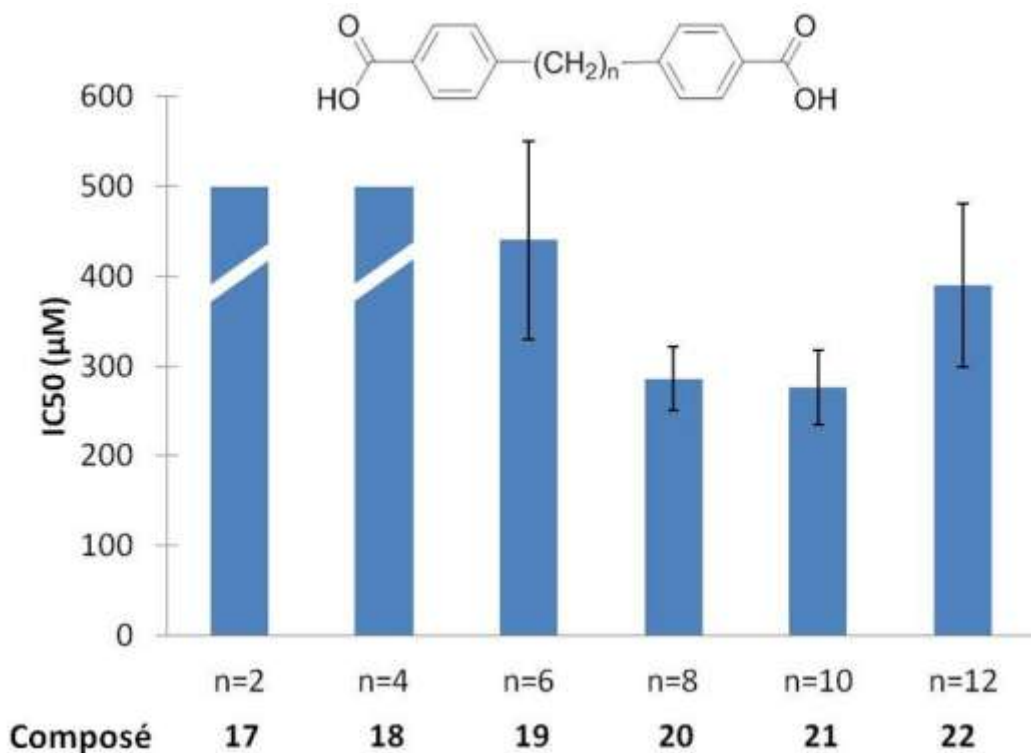


Figure 4.2 : Valeur des  $IC_{50}$  des composés **11** à **16**. Le  $IC_{50}$  du composé **16** n'a pas été établi (au dessus de 250  $\mu M$ ).

## Section 4.3 - Effet du groupement benzimidazole

Au chapitre 3, il fut déterminé que le composé **8a**, qui ne possède pas le fragment benzimidazole, ne provoquait pas l'inhibition de la DHFR R67 jusqu'à une concentration de 1000  $\mu M$ . Pour vérifier s'il était possible de combler le retrait du benzimidazole par l'élongation de la molécule, les composés **17** à **22** furent testés. L'effet de la longueur des

composés sur le  $IC_{50}$  est illustré à la figure 4.3. On observe un creux énergétique quand l'espaceur possède de 8 à 10 méthylènes. Par contre, les meilleurs  $IC_{50}$  se trouvent aux environs de 300  $\mu M$ , ce qui est plus faible que le  $IC_{50}$  des composés **12 à 14** ( $\approx 50 \mu M$ ). Ainsi, le retrait du benzimidazole ne peut pas être totalement comblé par l'élongation des molécules ce qui suggère que le benzimidazole établit des interactions favorables avec la DHFR R67.



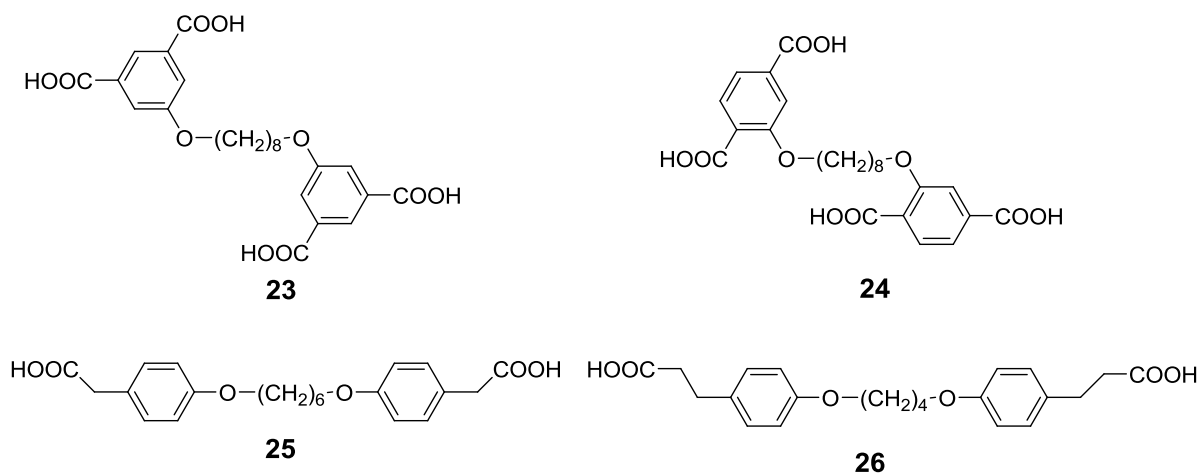
**Figure 4.3 :**  $IC_{50}$  des composés **17 à 22**. Les  $IC_{50}$  des composés **17** et **18** n'ont pas été établis (au dessus de 500  $\mu M$ )

## Section 4.4 - Design de nouveaux inhibiteurs

Dans le but de développer des nouveaux inhibiteurs plus puissants, une liste de propositions basées sur les informations obtenues durant les travaux des deux dernières années sera présentée. Une étude d'arrimage moléculaire viendra appuyer ces propositions. Ma contribution a été de proposer les modifications en tenant compte de leur activité biologique tandis que Dr. Delphine Forge (Université de Mons, Belgique) a proposé les modifications possibles en tenant compte de l'approche synthétique. Ces molécules seront synthétisés dans un futur proche.

### Section 4.4.1 - Modifications des extrémités

Au chapitre 3, on a démontré l'importance des fonctions carboxyliques. L'étude de modélisation suggère que ces carboxylates interagissent avec les Lys32 situé à l'entrée du site actif de la DHFR R67. La DHFR R67, étant un homotétramère, possède 4 résidus Lys32 (un pour chaque monomère) ainsi, il y a deux Lys32 à chacune des deux entrées de chaque côté du site actif. De plus, les résidus Lys33 se situent également aux entrées du site actif. En ajoutant des carboxylates additionnels sur les extrémités des molécules, il serait possible de former davantage d'interactions électrostatiques avec ces Lys 32 et 33, ce qui augmenterait la puissance des inhibiteurs (composé **23**). Modifier l'orientation de ces carboxylates pourrait optimiser les interactions électrostatiques avec les Lys32 et/ou Lys33 (composé **24**). Augmenter le degré de liberté des fonctions carboxyliques pourrait aussi favoriser les interactions électrostatiques (composé **25** et **26**).



**Figure 4.4 : Modifications proposés sur les extrémité des inhibiteurs dans le but de favoriser les interactions électrostatiques avec les Lys 32.**

Pour faciliter la synthèse, ces modifications seront apportées sur les inhibiteurs qui ne possèdent pas le fragment benzimidazole. Malgré que cette série d'inhibiteur était moins puissante, elle permettra de vérifier l'impact de ces modifications et de les introduire ultérieurement dans les inhibiteurs dotés de benzimidazole.

## Section 4.4.2 - Modifications des cycles

À la section 4.3, il fut proposé que le groupement benzimidazole forme des interactions favorables avec la DHFR R67, ce qui expliquait que les composés **20** et **21** sont moins puissants que les composés **9**, **10**, **12** à **14**. Dans le but de reproduire ces interactions favorables, on inclurait un atome d'azote dans le cycle du composé **20** (ce qui forme le composé **27**, figure 4.5). Si on arrive à combler l'absence du benzimidazole par l'ajout d'azote sur le cycle, on réduirait la masse moléculaire des inhibiteurs, ce qui augmenterait l'efficacité de liaison (LE).

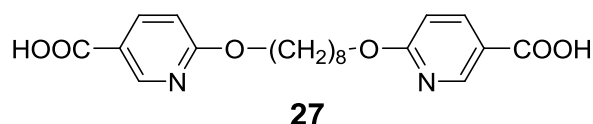
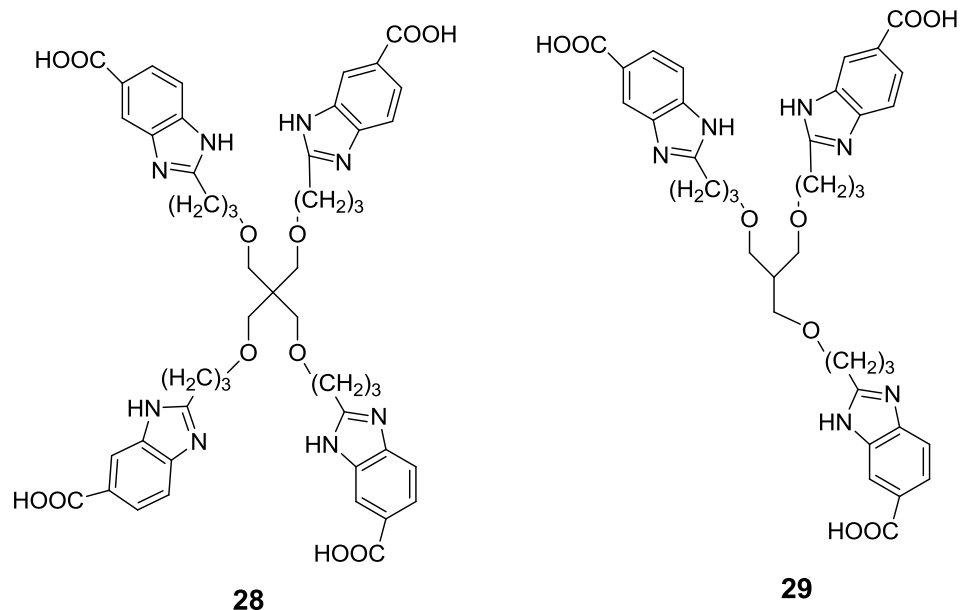


Figure 4.5 : Modification qui pourrait combler l'absence du fragment benzimidazole.

## Section 4.4.3 - Connexions de deux molécules d'inhibiteurs

Au chapitre 3, par le graphique de Hill, on a démontré que plus d'une molécule d'inhibiteur se lie dans le site actif de la DHFR R67. Ainsi, on propose de lier deux molécules d'inhibiteurs ensemble pour en former une qui serait plus puissante. La liaison d'une seule molécule au site actif, au lieu de deux, pourrait réduire la pénalité d'entropie de liaison (section 1.4.1). Ceci diminuerait l'énergie libre du système et ainsi, améliorerait le  $K_i$  (composé **28**). Dû au fait que le composé **28** est volumineux, le composé **29** sera synthétisé dans le but de réduire un éventuel encombrement stérique (figure 4.6). Les benzyles retrouvés sur les composés **8** et **11** à **16** ont été retirés dans le but de réduire la masse moléculaire de ces composés tout en incorporant une flexibilité dans cette molécule encombrée.



**Figure 4.6 : Connexions de bras de molécules inhibitrices de la DHFR R67 dans le but de réduire la pénalité d'entropie et améliorer la puissance des inhibiteurs.**

#### **Section 4.4.4 - Favoriser une conformation linéaire ou repliée**

Au chapitre 3, le modèle informatique obtenu avec les logiciels Moldock et Autodock Vina suggèrent que les inhibiteurs se replient en forme de U dans le site actif de la DHFR R67. Toutefois, le modèle avancé par le Conseil National de Recherche au Canada (CNRC) suggère que les inhibiteurs adoptent une conformation linéaire. Dans le but de vérifier ces modèles, les composés **30** et **31** seront testés (figure 4.7). Ces composés favorisent soit une conformation linéaire (**30**) ou en U (**31**). Lors de leur liaison sur la DHFR R67, ces composés perdraient moins d'entropie ce qui augmenterait leur puissance. Par exemple, si la vraie conformation est celle en U, le composé **30** perdrait moins d'entropie lors de sa liaison avec la DHFR R67 s'il adopte déjà la conformation en U en solution.

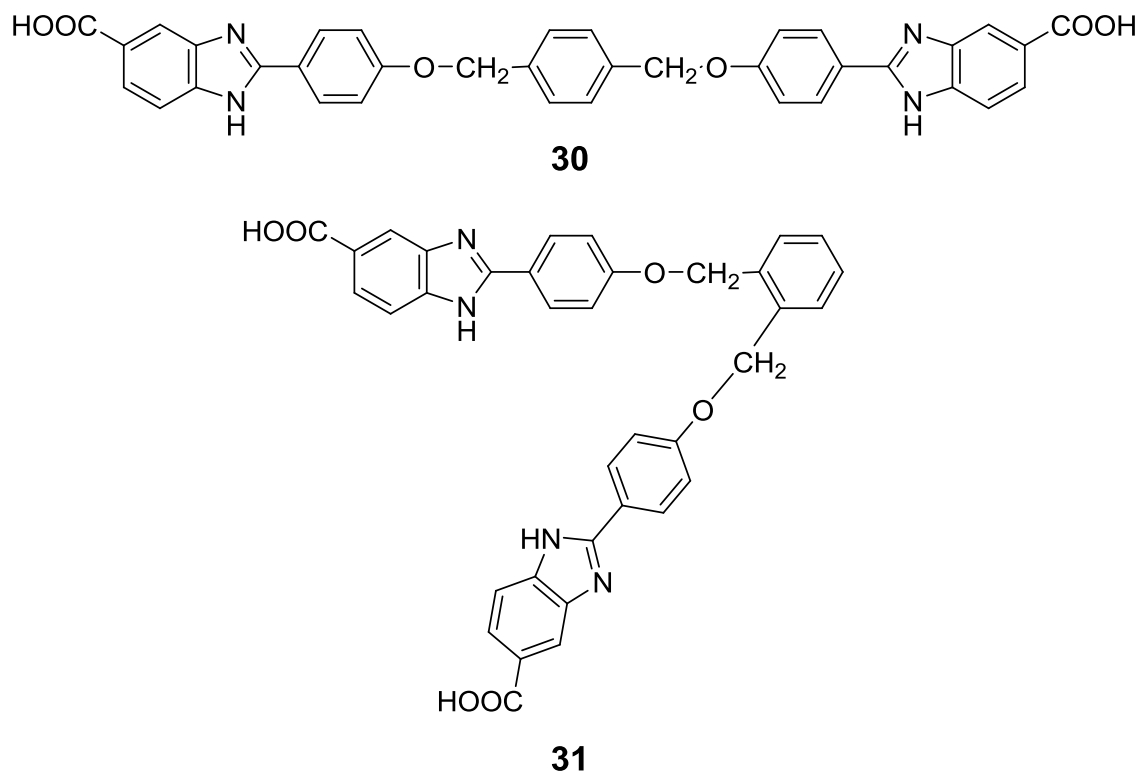


Figure 4.7 : Composés qui favorisent la conformation linéaire (30) ou repliée (31).

### Section 4.4.5 - Modifications des espaceurs

L'espaceur des composés **8** et **12** à **14** est simple et ne possède aucun hétéroatome outre les oxygènes qui le flanquent. L'étude de modélisation au chapitre 3 suggèrent que l'alcool central du composé **9** interagit fortement avec les résidus Ile68 (NH et CO de la chaîne principale). Dans la littérature, il est démontré que les résidus Ile68 interagissent (par leur groupement NH et CO sur leurs chaîne principale) avec les motifs amides qu'on retrouve sur le DHF et le NADPH (voir chapitre 1). De plus, l'absence d'amide sur le TMP serait une des raisons pour lesquelles le TMP ne peut se lier à la DHFR R67 (voir chapitre 1). Ainsi, en incluant des hétéroatomes dans l'espaceur, il serait possible d'améliorer la puissance des inhibiteurs. Le composé **33** possède deux amides qui pourraient interagir avec les Ile68. Pour le composé **34**, on a



ajouté deux alcools qui pourraient aussi interagir avec les Gln 67 et/ou Ile68. Finalement, l'ajout d'un azote au centre de l'espaceur (composé **35**) pourrait former des pont H avec les Gln67 et/ou Ile68.

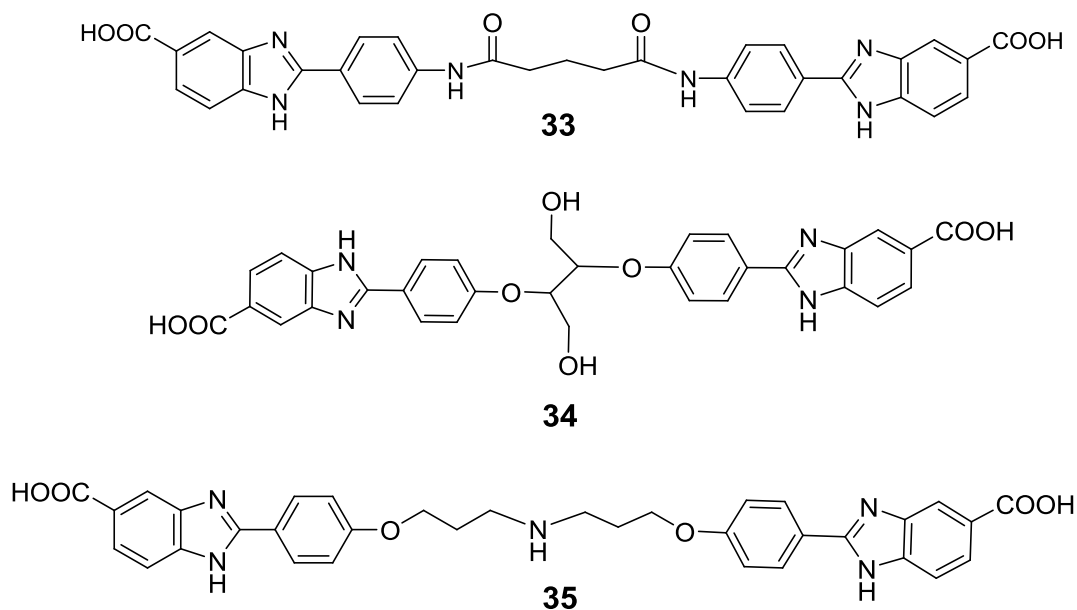


Figure 4.8 : Modification de l'espaceur

## Section 4.5 - Arrimage moléculaire des composés **23** à **34**

Dans le but de vérifier l'effet des modifications proposées à la section précédente, une étude d'arrimage moléculaire a été faite. Le tableau 4.2 fait état des scores des composés **23** à **35**. Les composés ont été placés en ordre croissant de score. Le score du composé **8** et du NADPH ont été inclus à titre de contrôle positifs de molécules qui se lient à la DHFR R67.

**Tableau 4.2 : Arrimage moléculaire des composés 23 à 24**

Composé	Moldock Score
<b>NADPH</b>	<b>-190</b>
<b>8</b>	<b>-190</b>
28	-236
33	-224
31	-220
30	-217
34	-200
29	-190
35	-185
23	-130
25	-129
24	-128
26	-122
27	-119

Les composés **28**, **30**, **31**, **33** et **34** ont tous obtenu un meilleur score que le NADPH et le composé **8**, ce qui suggère qu'ils devraient se lier plus fortement que ces derniers. Le composé **28** a obtenu le meilleur score de tous, soit de -236. Ainsi, connecter deux molécules d'inhibiteurs **8** semble être favorable. Des interactions semblables à celle du composé **8** (chapitre 3) avec les résidus Lys32, Ala36, Tyr 46, Gln67, Ile68 et Tyr69 sont observées dans le modèle. Toutefois, elles sont plus nombreuses du fait que la molécule est plus grosse, ce qui explique le meilleur score. Les modifications sur l'espaceur semblent aussi être favorables comme le suggère le score pour les composés **33** et **34**. L'ajout d'amides sur l'espaceur est une modification favorable, ce qui est cohérent avec la littérature (voir chapitre 1). Dans le modèle du composé **33**, des interactions entre l'espaceur et des Gln67 et Ile68 sont observées. Pour ce qui est du composé **34**, des interactions similaires sont également observées dans le modèle. L'ajout d'un azote sur l'espaceur (composé **35**) n'a pas eu un grand changement, le score étant similaire au composé **8** et au NADPH. Dans le modèle du composé **35**, l'azote ne forme pas d'interactions avec des résidus.

Les composés **30** et **31**, dont la conception vise à favoriser respectivement l'adoption de la conformation linéaire ou en U, ont obtenu un meilleur score que le NADPH et le composé **8**. Par contre, dans les deux cas, les composés ont adopté une conformation en U dans le site actif de la DHFR R67 lors de l'arrimage informatique. Le logiciel ne peut modéliser la trajectoire du composé lors de son entrée dans le site actif, donc il se pourrait que l'arrimage moléculaire ne soit pas une technique adéquate pour vérifier l'importance d'une conformation linéaire ou repliée. Par contre, l'ajout d'un benzyle sur l'espaceur semble être favorable puisque le score pour les deux composés est meilleur que celui du NADPH et du composé **8**. Des interactions avec les Lys32, Tyr 46, Thr51 sont observés dans le modèle des composés **30** et **31**.

Finalement, les composés **23** à **27** ont tous obtenu un score nettement plus faible que le NADPH et le composé **8** ce qui suggère que le fragment benzimidazole est important. L'ajout de fonctions carboxyliques n'a pas eu un grand effet. Dans les modèles, on retrouve rarement plus que deux interactions électrostatiques à la fois. Par contre, dans la majorité des cas, les composés sont repliés en U. Tel qu'expliqué au chapitre 3, le logiciel pourrait être biaisé et replier les molécules en U pour remplir l'espace et maximiser les contacts avec le site actif de la DHFR R67. Il se pourrait qu'en repliant en U les molécules, trop de charges négatives se retrouvent dans un espace restreint ce qui provoque de la répulsion et augmente le score.

Selon ces résultats de modélisation, une priorité de synthèse sera établie avec nos collaborateurs (Dr. Delphine Forge, Dr. Jean-Jacques Vanden Eynde, Université de Mons, Belgique). Les composés seront évalués expérimentalement tel que décrit dans les chapitres précédents.

## Chapitre 5 - Criblage par arrimage moléculaire

Parallèlement au design à base de fragments, deux campagnes de criblages par arrimage moléculaire ont été entreprises afin de découvrir de nouveaux inhibiteurs de la DHFR R67. Le premier criblage a été accompli avec le logiciel Moldock (Molegro computational drug discovery) sur la librairie virtuelle de Maybridge (version 4.0.0, année 2010). Le deuxième criblage a été fait en collaboration avec la plate-forme de modélisation de l'institution de Recherche en Biotechnologie du Conseil National de Recherche du Canada (IRB-CNRC) sur leur librairie maison qui est dérivée de la librairie virtuelle ZINC.

### Section 5.1 - Criblage de la librairie virtuelle Maybridge

Dû à la puissance de calcul limitée disponible au laboratoire Pelletier, nous avons choisi une librairie de taille plus modeste. Comparativement à d'autres librairies comme celle de Chembridge ou Enamine, qui contiennent chacune des millions de composés, celle de Maybridge en possède environ 80 000 (été 2010). Choisir une librairie plus petite permettait de la cribler efficacement dans un temps raisonnable. Malgré qu'elle soit plus petite, elle est diversifiée ce qui permet de cribler efficacement l'espace chimique. Certaines librairies plus volumineuses comprennent beaucoup de composés analogues, ce qui peut diminuer l'intérêt de les cribler. En effet, il est rare que les analogues rapprochés donnent des scores qui diffèrent grandement. Donc, il est plus efficace de cribler une petite banque diversifiée que de cribler une très grande banque ayant une petite diversité.

La librairie de Maybridge est composée de molécules qui respectent la règle de 5 de Lipinski et un bon profil ADME est prédit pour les composés de cette banque. La règle de Lipinski provient d'une analyse empirique qui a été faite sur les médicaments. En règle générale, les médicaments possèdent une masse molaire en dessous de 500 g/mol, possèdent moins de 5 donneurs de pont H et moins de 10 accepteurs de pont H et leurs logP (solubilité) est en dessous de 5.

Ainsi, des molécules qui respectent ces critères ressembleraient déjà à un médicament. Par contre, cette règle n'est pas absolue, car certains médicaments ne respectent pas cette règle. Par exemple, le Taxol possède une masse molaire d'environ 850 g/mol, 15 accepteurs de pont H et 4 donneurs de pont H. Suivre à la lettre la règle de Lipinski pourrait empêcher de découvrir des molécules pour certaines cibles. Par contre, dans le cas de la DHFR R67, il s'agit d'une nouvelle cible et commencer par cribler des molécules qui respectent la règle de Lipinski est approprié. Si le criblage est un succès, les composés possèderaient déjà des propriétés médicamenteuses, ce qui pourrait faciliter les étapes d'optimisations.

### **Section 5.1.1 - Choix du logiciel Moldock**

Comme il a été mentionné au chapitre 1, il existe plusieurs logiciels qui fonctionnent chacun avec des algorithmes différents. Moldock est un logiciel récemment développé qui a démontré un meilleur taux de succès que les autres logiciels fréquemment utilisés dans le domaine tels que Glide, Surflex, FlexX et GOLD<sup>31</sup>. Moldock utilise un algorithme génétique (chapitre 1). La fonction de score de Moldock est une fonction empirique (chapitre 1). Bien que ce type de fonction ne permette pas de prédire une énergie de liaison absolue, elle donne une évaluation relative de l'énergie de liaison, ce qui permet de choisir les composés ayant la plus grande probabilité d'interagir avec la cible. De plus, ce type de fonction ne nécessite pas une grande puissance de calcul, ce qui permet de cribler la cible plus rapidement.

La fonction de score de Moldock se démarque par sa capacité à mieux évaluer les ponts H ainsi que les interactions électrostatiques<sup>31</sup>. Puisque les interactions électrostatiques contribuent de façon importante dans la liaison du DHF et du NADPH au site actif de la DHFR R67, le logiciel Moldock est un choix approprié pour entreprendre le criblage virtuel de la librairie Maybridge sur la DHFR R67.

## Section 5.1.2 - Résultats du criblage virtuel

Les résultats du virtuel des 80 000 composés de la librairie Maybridge sont présentés en ordre du meilleur score jusqu'au score le plus faible (tableau 5.1; 60 meilleurs composés).

Tableau 5.1 : Soixante composés ayant obtenu le meilleur score suite au criblage de la librairie Maybridge

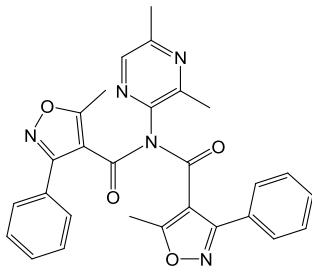
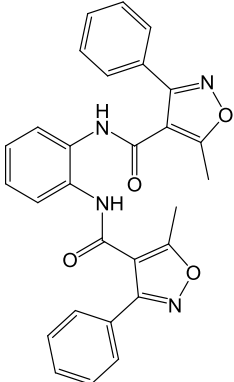
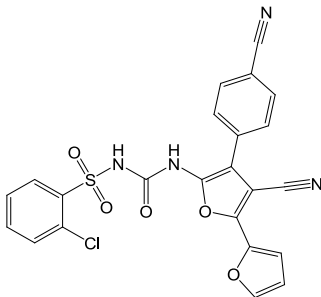
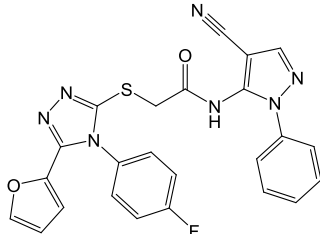
Code <sup>a</sup>	Moldock score	Code	Moldock score	Code	Moldock score
ZINC12340441	-193.69	ZINC01048606	-172.97	ZINC01047137	-168.30
ZINC08616652	-193.48	ZINC03135257	-172.92	ZINC08383004	-168.11
ZINC08628199	-183.76	ZINC08445450	-172.74	ZINC01035409	-167.86
ZINC08636932	-179.67	ZINC01033924	-172.63	ZINC12339424	-167.62
ZINC12340610	-179.60	ZINC08636658	-172.48	ZINC12340714	-167.55
ZINC12339423	-179.34	ZINC12339800	-172.41	ZINC12367644	-167.50
ZINC01026810	-178.69	ZINC12339358	-172.38	ZINC12369525	-167.49
ZINC01033926	-178.16	ZINC19880224	-172.32	ZINC12340719	-167.43
ZINC12340664	-177.77	ZINC08648867	-172.04	ZINC12340566	-167.38
ZINC13465491	-177.47	ZINC12339359	-171.43	ZINC08628199	-167.24
ZINC01033923	-177.23	ZINC08618167	-171.08	ZINC03119964	-166.89
ZINC01036093	-177.06	ZINC01047585	-171.01	ZINC01045842	-166.50
ZINC08628198	-175.51	ZINC01034256	-170.98	ZINC08637021	-166.47
ZINC12339800	-175.17	ZINC01035379	-170.95	ZINC01034038	-166.44
ZINC08647767	-175.09	ZINC12339067	-170.57	ZINC01034696	-166.43
ZINC01047135	-174.83	ZINC01035443	-170.31	ZINC01036504	-166.39
ZINC12339359	-173.80	ZINC01030029	-169.77	ZINC08636783	-166.30
ZINC12340726	-173.32	ZINC08926016	-169.09	ZINC12339129	-166.25
ZINC08647769	-173.27	ZINC01036506	-168.70	ZINC08647872	-166.06
ZINC12370086	-173.22	ZINC12340607	-168.55	ZINC12370084	-166.01

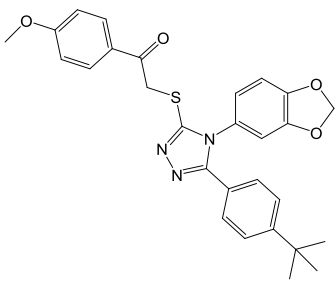
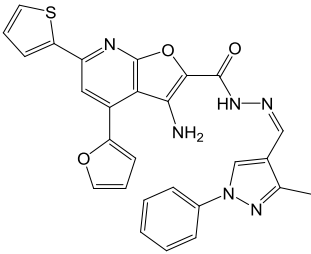
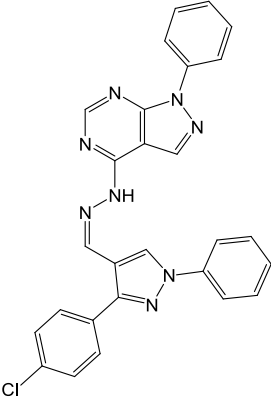
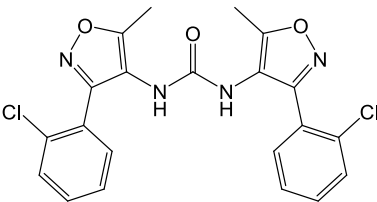
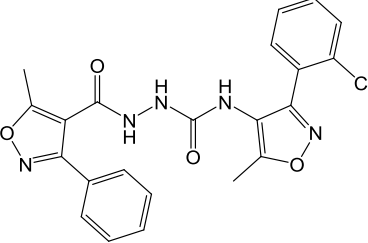
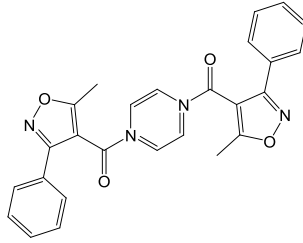
<sup>a</sup>Les structures correspondant aux codes se trouvent sur le site [www.ZINC.docking.org](http://www.ZINC.docking.org)

Le meilleur composé a obtenu un score de -193.69 et le 60ième de -166.01. Les composés situés aux derniers rangs avaient des scores d'environ -100. On constate donc que Moldock a été en mesure de classer un nombre restreint de composés, parmi la banque de 80 000, ayant un meilleur potentiel de liaison à la DHFR R67.

En vue de procéder aux tests *in vitro* une analyse manuelle des 60 meilleurs composés a été entreprise. Les poses et les structures de ces 60 composés ont été étudiées afin de choisir des composés ayant des interactions prédites cohérentes avec la littérature, une solubilité prédite adéquate ainsi qu'une structure différente afin de cribler le plus grand espace chimique possible. Le tableau 5.2 illustre les structures des dix composés choisis pour des tests d'inhibition.

**Tableau 5.2 : Composé choisi pour les tests *in vitro***

Code	Structure	Code	Structure
ZINC12340441		ZINC01033926	
ZINC08616652		ZINC01036093	

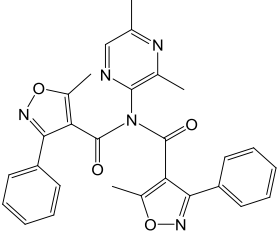
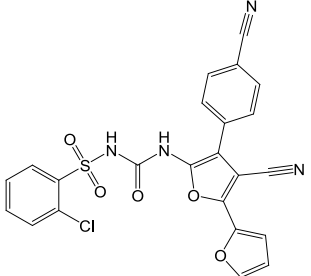
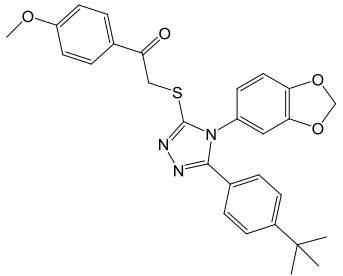
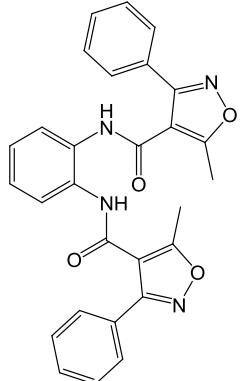
ZINC08628199		ZINC12340726	
ZINC12340610		ZINC01048606	
ZINC01026810		ZINC01045842	

### Section 5.1.3 - Test *in vitro* de dix des composés Maybridge identifiés *in silico*

Les dix composés du tableau 5.2 ont été testés *in vitro* en tant qu'inhibiteur de la DHFR R67. Parmi ces dix composés, quatre ont inhibé la DHFR R67 dans l'ordre du micromolaire (tableau 5.3).



**Tableau 5.3 : Composés Maybridge offrant une inhibition de la DHFR R67**

Code	Structure	Moldock score	IC50 (μM)
ZINC12340441		-193.69	300 ± 40
ZINC08616652		-193.48	160 ± 50
ZINC08628199		-183.76	130 ± 50
ZINC01033926		-178.16	120 ± 30

Il est intéressant de noter que les composés ZINC12340441 et ZINC01033926 sont symétriques et possèdent tous les deux des groupements phenylisoxazocarboxamide. Les deux autres composés, le ZINC08616652 et le ZINC08628199, seraient les premiers inhibiteurs micromolaires asymétriques découverts à ce jour. Il est important de noter que la solubilité de ces composés était très

faible et il a été difficile de reproduire ces résultats. Il sera important de démontrer la liaison de ces composés à la DHFR R67 selon une technique alternative. La calorimétrie isotherme à titration (ICT) pourrait être une technique appropriée pour vérifier si la chute d'activité observée provient de la liaison des molécules à la DHFR R67. L'ITC permettrait aussi de déterminer les  $K_i$  ainsi que la stœchiométrie.

## **Section 5.2 - Criblage de la librairie du CNRC**

En collaboration avec la plate-forme de modélisation de l'IRB-CNRC (Dr. Enrico Purisima, Dr. Train Sulea et Dr. Hervé Hogues), un 2<sup>ème</sup> criblage virtuel fut effectué sur la DHFR R67 avec des algorithmes d'arrimage moléculaire développés par le CNRC<sup>42</sup>. La librairie maison du CNRC est dérivée de la banque virtuelle ZINC qui contient près de 14 millions de composés disponibles commercialement. Le CNRC a conçu sa banque en filtrant ces 14 millions de composés pour rejeter tous les composés qui ne respectaient pas la règle de Lipinski ainsi que les composés qui possédaient plus de 7 rotations libres et/ou des charges, car de tels composés sont difficiles à modéliser. Finalement, ils ont aussi rejeté les composés trop similaires afin de diminuer la taille de leur banque tout en maintenant sa diversité. Ainsi, la banque virtuelle du CNRC possède environ 5.6 millions de composés. Ce criblage a été fait en deux étapes : une première ronde de criblage a été faite sur la banque complète avec l'algorithme Wilma qui est spécialisé dans la prédiction des poses. Par la suite, une 2<sup>ème</sup> ronde de criblage a été entreprise sur les meilleurs composés de la première ronde avec l'algorithme SIE qui se spécialise dans la prédiction de l'affinité. La plate-forme de modélisation du CNRC possède la puissance de calcul nécessaire pour cribler cette banque en moins de deux semaines.

Le criblage virtuel du CNRC a été entrepris en utilisant deux conformations de la DHFR R67, soit la conformation apoenzyme et holoenzyme. Dans la forme holoenzyme, les Gln67 de la DHFR R67 sont orientées vers la paroi du tunnel du site actif (figure 5.1A). En contraste, dans la conformation apoenzyme, les Gln67 sont orientés vers le centre du site actif (Figure 5.1B). Un inhibiteur peut avoir d'avantage

d'affinité envers une des deux formes. En criblant la librairie sur ces deux formes de la DHFR R67, on augmente le taux de succès.

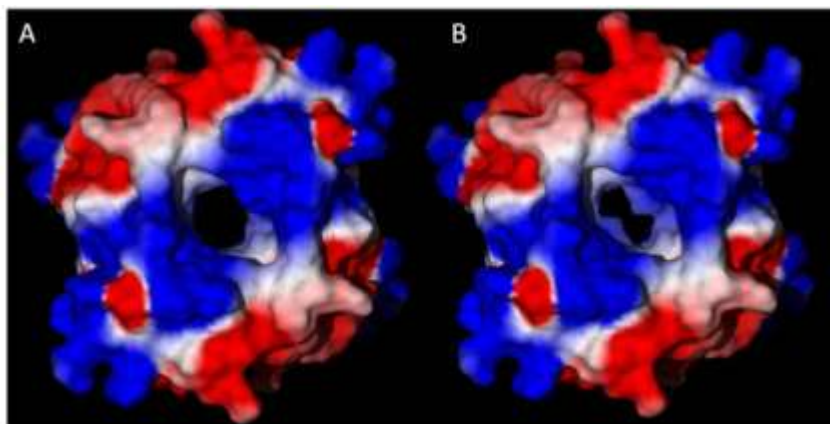


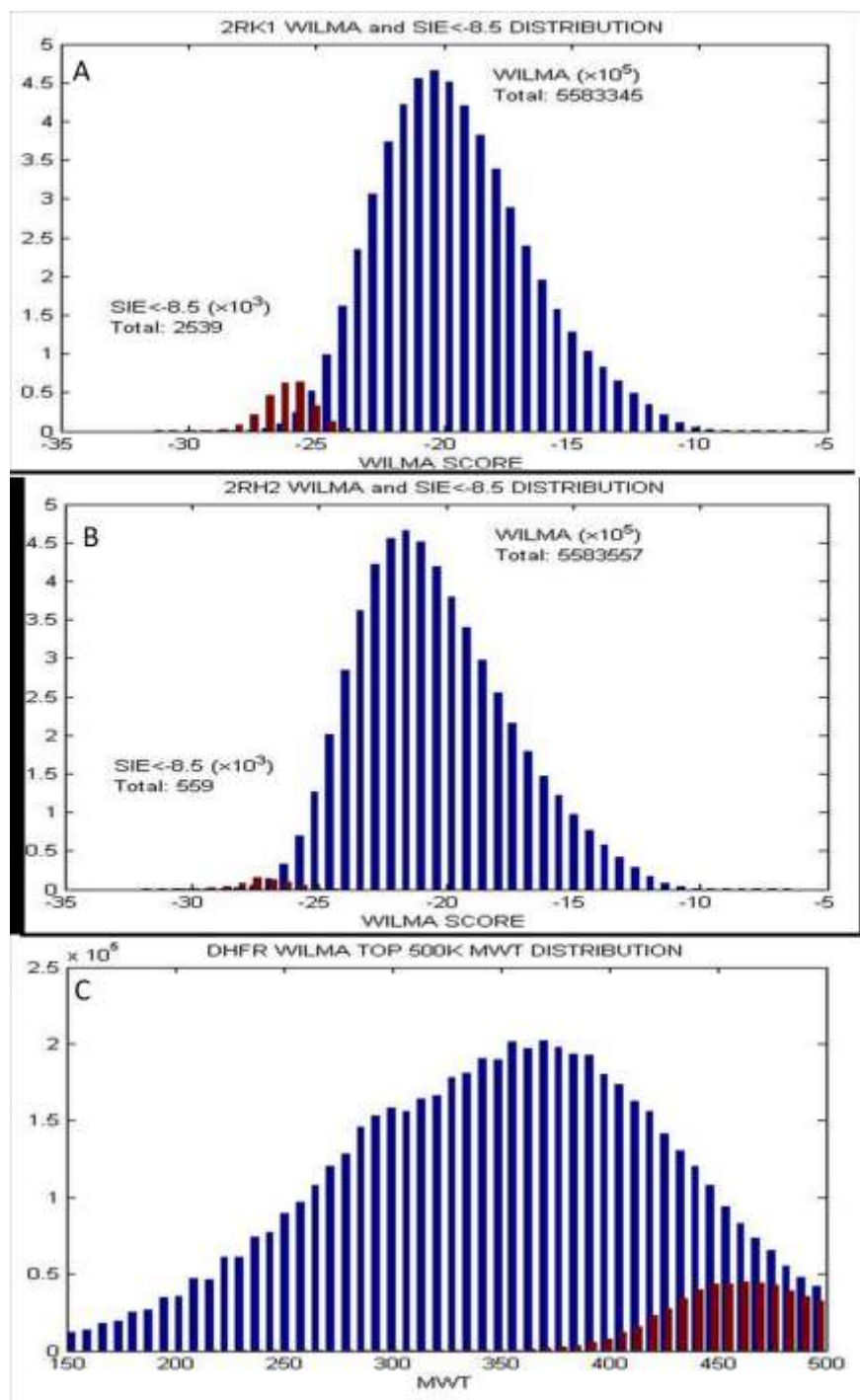
Figure 5.1: Forme holoenzyme(A) et apoenzyme (B) de la DHFR R67. Dans la forme holoenzyme, les Gln67 se trouvent repliées contre la paroi du site actif. Dans la forme apoenzyme, les Gln67 occupent le site actif.

### Section 5.2.1 - Résultats du criblage de la librairie du CNRC

Sur la figure 5.2 (holoenzyme 5.2A, apoenzyme 5.2B), en bleu, on peut observer la distribution des scores de la première ronde entreprise avec l'algorithme Wilma. Les scores varient de -5 à -34. Le score de référence (NADPH) est de -32.6. Ainsi, selon cette prédiction, aucun composé dans la banque du CNRC ne se lierait significativement mieux que le NADPH.

Sur la figure 5.2A et 5.2B, en rouge, on peut observer la distribution des scores de la 2<sup>ème</sup> ronde, où les meilleurs composés identifiés lors de la première ronde ont été évalués par l'algorithme SIE<sup>42</sup>. Le score de référence (NADPH) est de -8.82. Pour les 500 meilleurs composés du criblage de la forme apoenzyme, les scores SIE varient de -9.26 à -8.68. Les scores SIE pour les 500 meilleurs composés de la forme holoenzyme varient de -9.06 à -8.51. Ainsi, les scores sont très similaires et n'offre pas une discrimination d'un nombre restreint de composés qui seraient significativement meilleurs. Ce modèle prédit aussi que ces composés se lieraient aussi bien que le NADPH, au mieux.

La figure 5.2 C démontre que les scores sont biaisés par la masse molaire des composés. En effet, on constate que plus un composé a une masse molaire élevée, plus son score est bon. Plus de 90 % des composés ayant une masse molaire de 500 g/mol étaient parmi les meilleurs tandis qu'à peine 5 % des composés ayant une masse molaire de 400 g/mol se sont retrouvés parmi les meilleurs. Le site actif de la DHFR R67 est principalement hydrophobe; il se pourrait que les interactions de Van der Waals jouent un rôle important, ce qui expliquerait que plus les composés sont gros, plus ils se lient fortement. Ainsi, ces résultats suggèrent qu'un puissant inhibiteur de la DHFR R67 pourrait ne pas suivre la règle de Lipinski et qu'il pourrait être problématique de se limiter à des composés qui respectent cette règle.



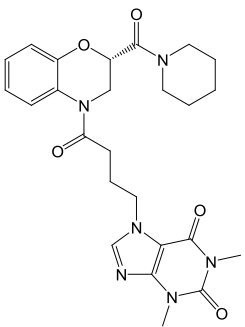
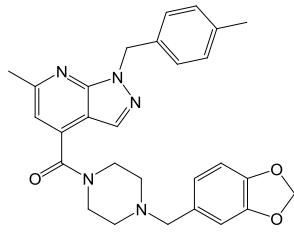
**Figure 5.2: Distribution selon les scores (A et B) et la masse molaire (C) suite au criblage virtuel de la librairie du CNRC.** (A) Distribution des scores obtenus lors du criblage de la forme holoenzyme. En bleu, on peut voir la distribution du score obtenu selon l'algorithme Wilma. En rouge, la distribution des scores selon l'algorithme SIE. (B) Distribution des scores obtenu lors du criblage de la forme apoenzyme. En bleu, la distribution du score obtenu selon l'algorithme Wilma. En rouge, la distribution des scores selon l'algorithme SIE. (C) Distribution des 500 000 meilleurs composés selon leur masse molaire. En bleu, la distribution des masses molaires des composés de la librairie virtuelle complète du CNRC ( $5.6 \times 10^6$  composés). En rouge, la distribution de la masse molaire 500 000 meilleurs composés. Plus de 90 % des 500 000 composés retenus ont une masse moléculaire au dessus de 400g/mol.

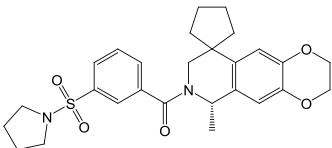
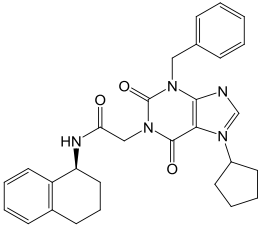
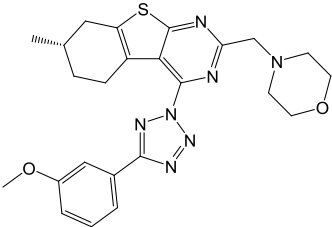
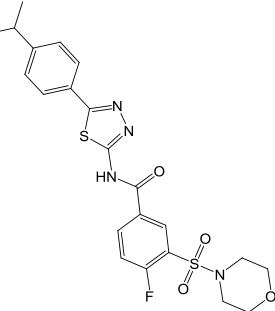
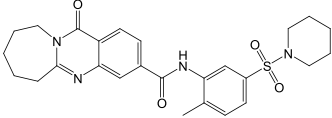
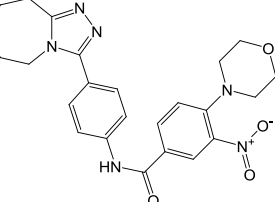
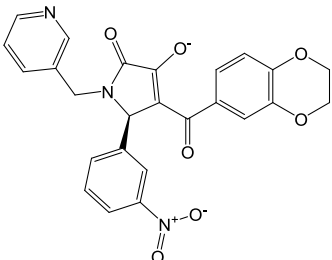
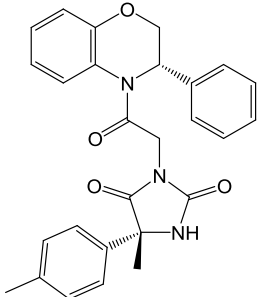
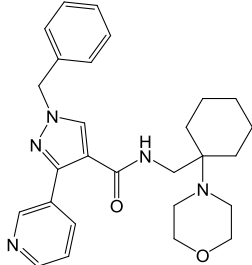
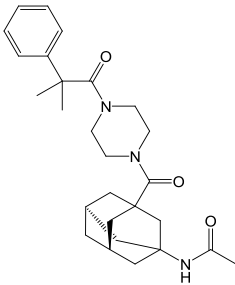
## Section 5.2.2 - Test *in vitro* des composés du CNRC

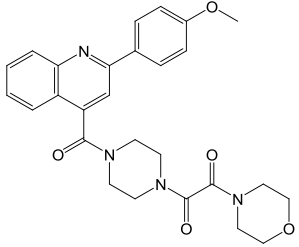
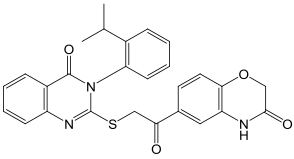
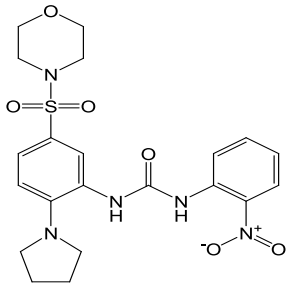
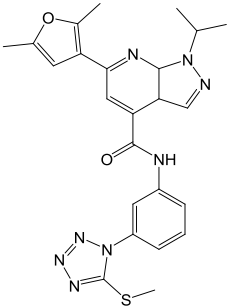
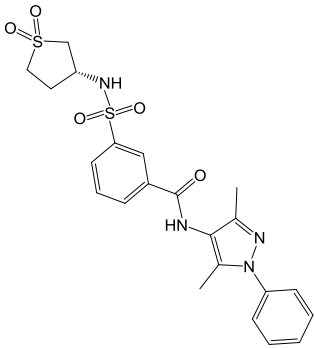
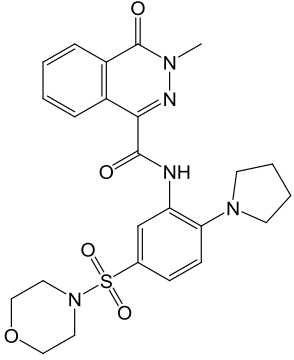
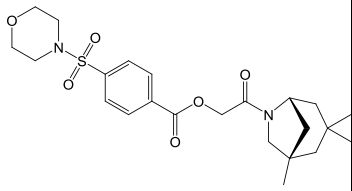
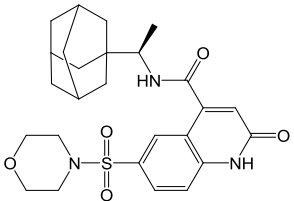
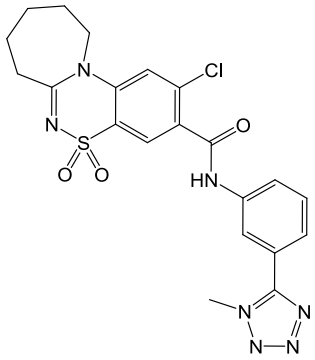
Malgré les biais révélés par le criblage virtuel entrepris par le CNRC, on a procédé à une sélection de 21 composés pour entreprendre des tests d'inhibitions *in vitro*. Il a été prédit que les meilleurs composés pourraient se lier aussi fortement que le NADPH ( $K_M$  de 1.6  $\mu$ M). Les composés dans la banque du CNRC ne possèdent pas de charges négatives. Il a été démontré que les charges négatives sur le NADPH<sup>3</sup>, DHF<sup>3</sup> et l'inhibiteur **8**<sup>4</sup> (ajouter Ref) étaient importantes dans la liaison de ces molécules avec la DHFR R67. Il serait intéressant de découvrir des inhibiteurs possédant une affinité dans les environs de 2  $\mu$ M qui ne possèdent pas de charges négatives. Ajouter de telles charges à ces molécules par la suite pourrait améliorer grandement l'affinité de ces inhibiteurs ainsi que leur solubilité.

Les résultats du criblage virtuel du CNRC ont été divisés en 3 séries : les meilleurs composés provenant du criblage de la forme apoenzyme, du criblage de la forme holoenzyme et ceux ayant obtenu un bon score sur les deux formes. Sept composés de chaque série de résultats furent testés *in vitro*. Dû aux scores similaires obtenus, nous avons choisi des composés représentant la plus grande diversité chimique possible (tableau 5.4).

**Tableau 5.4 : Composés provenant du criblage virtuel du CNRC choisis pour les tests *in vitro*.**

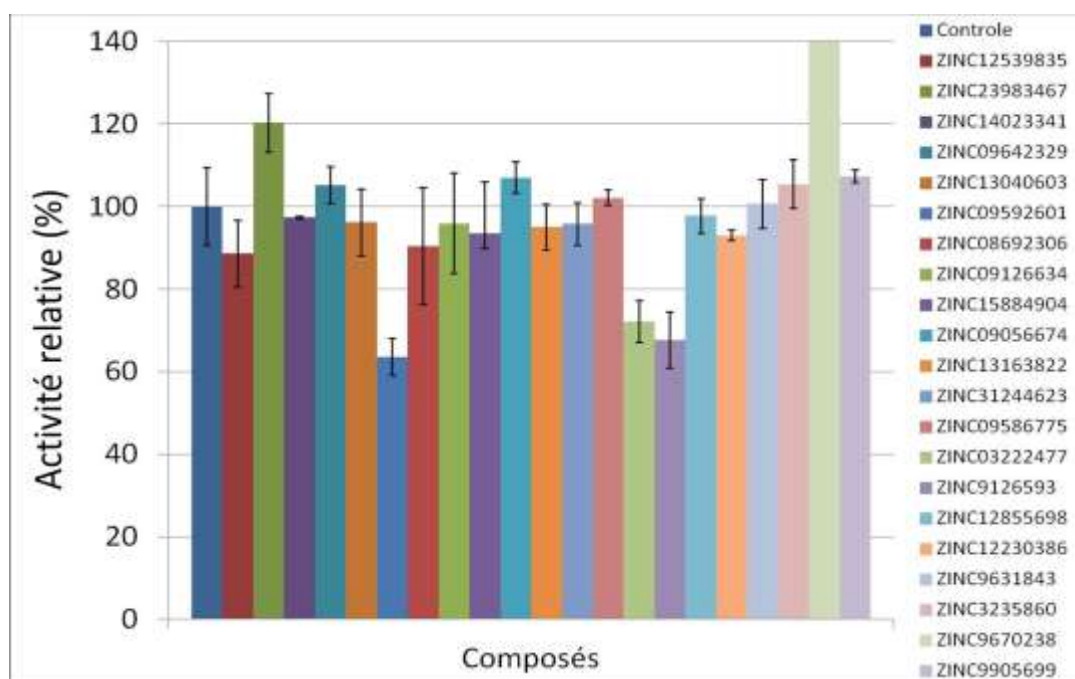
Code	Structure	Code	Structure
ZINC12539835		ZINC23983467	

ZINC14023341		ZINC09642329	
ZINC13040603		ZINC09592601	
ZINC08692306		ZINC09126634	
ZINC15884904		ZINC09056674	
ZINC13163822		ZINC31244623	

ZINC09586775		ZINC03222477	
ZINC9126593		ZINC12855698	
ZINC12230386		ZINC9631843	
ZINC3235860		ZINC9670238	
ZINC9905699			



Le criblage de ces composés a été effectué à une concentration à 100  $\mu\text{M}$  (figure 5.3). Sur les 21 composés testés, 3 composés ont démontré une activité inhibitrice contre la DHFR R67 à une concentration de 100  $\mu\text{M}$ : les composés ZINC09592601, ZINC03222477 et ZINC9126593 ont diminué l'activité de la DHFR R67 à environ 60 – 70% de l'activité non-inhibée.



**Figure 5.3 : Tests *in vitro* des 21 composés provenant du criblage virtuel du CNRC.** L'activité résiduelle de la DHFR R67 a été mesurée en présence de 100  $\mu\text{M}$  de chaque composé puis a été normalisée selon le contrôle positif. Le composé ZINC9670238 n'était pas soluble à 100  $\mu\text{M}$  ce qui a donné lieu à des résultats erratiques.

La figure 5.4, représente les structures des 3 inhibiteurs découverts. Le composé ZINC09592601 provient de la série apoenzyme tandis que les composés ZINC03222477 et ZINC9126593 proviennent de la série de composés ayant obtenu un bon score sur les deux conformations de la DHFR R67. Les 3 composés possèdent des groupements chimiques assez différents et sont asymétriques. Il est intéressant de remarquer qu'ils possèdent de nombreux cycles aromatiques et sont plutôt planaires; il y a une prédominance de carbones  $\text{sp}^2$  par rapport aux carbones  $\text{sp}^3$ . Le composé ZINC03222477 possède deux amides qui pourraient interagir avec la chaîne principale des Ile68. Le composé ZINC9126593 possède un groupement nitro fortement polarisé ce qui permettrait des interactions électrostatiques avec les Lys32.



**Figure 5.4: Structure des 3 inhibiteurs provenant du criblage virtuel du CNRC.**

Les résultats d'inhibition n'ayant été obtenus qu'à une concentration, il sera important d'approfondir l'étude de ces composés en faisant un test dose-réponse. Par contre, la faible solubilité de ces composés au-dessus de 100  $\mu\text{M}$  pourrait rendre difficile d'exécuter de tels tests. Tester ces composés selon une méthode alternative, comme l'ITC, serait important afin de démontrer hors de tout doute si ces molécules se lient à la DHFR R67.

## Références du chapitre 5

1. Thomsen, R.; Christensen, M. H., MolDock: a new technique for high-accuracy molecular docking. *J Med Chem* **2006**, *49* (11), 3315-21.
2. Sulea, T.; Hogues, H.; Purisima, E. O., Exhaustive search and solvated interaction energy (SIE) for virtual screening and affinity prediction. *J Comput Aided Mol Des* **2011**.
3. Krahn, J. M.; Jackson, M. R.; DeRose, E. F.; Howell, E. E.; London, R. E., Crystal structure of a type II dihydrofolate reductase catalytic ternary complex. *Biochemistry* **2007**, *46* (51), 14878-88.
4. Bastien, D.; Ebert, M. C.; Forge, D.; Toulouse, J.; Kadnikova, N.; Perron, F.; Mayence, A.; Huang, T. L.; Vanden Eynde, J. J.; Pelletier, J. N., Fragment-based design of symmetrical bis-benzimidazoles as selective inhibitors of the trimethoprim-resistant, type II R67 dihydrofolate reductase. *J Med Chem* **2012**, *55* (7), 3182-92.

## Chapitre 6 - Conclusion

Lors de ce projet de recherche, un essai d'activité automatisé fut développé pour cribler des composés en tant qu'inhibiteurs potentiels de la DHFR R67. Ce test d'activité a permis d'appliquer une approche de design à base de fragments permettant l'identification d'une série de molécules inhibant faiblement la DHFR R67. Ces molécules ont servi à construire des molécules plus complexes et plus puissantes dont la symétrie est le reflet du site actif de l'enzyme cible. L'étude d'analogues de ces inhibiteurs jumelé à des études de modélisation a permis de construire un modèle qui suggère que ces inhibiteurs se replient en forme de "U" dans le site actif de la DHFR R67. Ceci impliquerait des interactions d'empilements intramoléculaires. De plus, le modèle suggère que les résidus Lys32 formeraient des interactions électrostatiques avec les carboxylates présents sur les inhibiteurs.

Parallèlement, deux criblages virtuels ont été entrepris, permettant d'identifier sept inhibiteurs micromolaires. Par contre, la faible solubilité de ces composés n'a pas permis de confirmer hors de tout doute que ces composés étaient réellement des inhibiteurs. Afin de démontrer que la diminution de l'activité de la DHFR R67 provient de l'interaction directe des composés avec la DHFR R67, il sera important de développer une méthode complémentaire au test d'activité. À la section 1.3.2, plusieurs approches pouvant servir à détecter l'interaction entre un composé et une protéine furent présentées. Dû aux particularités de la DHFR R67 (nature tétramérique et symétrie 222), plusieurs de ses techniques (SPR, spectrométrie de masse, RMN) seraient difficiles, voir impossibles, à appliquer. Par contre, deux techniques, soit le titrage calorimétrique isotherme (ITC) et la cristallographie pourraient être envisagées.

L'ITC est une technique qui fut exploitée dans plusieurs études de la DHFR R67<sup>43,44,45</sup>. Lors de ces études, l'ITC a servi à déterminer des constantes cinétiques telles le  $K_D$  et le  $K_M$  pour le DHF et le NADPH. Bien que la caractérisation d'inhibiteurs sur la DHFR R67 n'a jamais été entreprise par l'ITC, il serait sûrement possible de l'appliquer afin de déterminer l'affinité des inhibiteurs avec la DHFR R67. En plus de servir de méthode complémentaire au test d'activité de la DHFR R67, l'ITC permettrait de déterminer

d'autres paramètres ne pouvant pas être déterminés avec le test d'activité tel que le  $K_D$ ,  $\Delta H$ ,  $\Delta S$ ,  $\Delta G$  de liaison et la stoechiométrie de liaison.

La cristallographie pourrait aussi être envisagée comme technique complémentaire. Tel que démontré, la structure de la DHFR R67 a déjà été résolue. Cette technique pourrait servir à confirmer que les composés interagissent avec la DHFR R67. De plus, il serait possible d'observer comment les composés interagissent avec la protéine, et du même fait, déterminer la stoechiométrie, ce qui ultimement donnerait d'importantes informations dans le développement d'inhibiteurs plus puissants. L'évaluation des composés mentionnés au chapitre 4 pourrait permettre d'identifier un inhibiteur plus puissant et d'obtenir davantage d'informations au niveau de la relation structure-activité (SAR) de cette série d'inhibiteurs sur la DHFR R67.

Suite à l'identification de composés plus puissants, des études pharmacologiques *in vitro* pourraient être envisagées. Il serait possible de déterminer un profil pharmacocinétique partiel afin de prédire si un composé aurait une bonne biodisponibilité. Des facteurs tels que l'absorption et le métabolisme peuvent influencer la biodisponibilité. L'absorption dans le corps humain peut être influencée par plusieurs facteurs tels que la solubilité, le  $pK_a$ , la lipophilicité et la perméabilité membranaire. La perméabilité membranaire peut être déterminée *in vitro* par des expériences d'essais de transports où l'on détermine si le composé d'intérêt diffuse à travers les membranes et/ou s'il est un substrat pour les pompes d'efflux et d'influx. Le métabolisme de phase I et II peut modifier les composés et ainsi réduire la biodisponibilité. Le métabolisme de phase I consiste en l'oxydation des composés par des enzymes tels que les cytochromes P450 qui se retrouvent principalement dans le foie. Le métabolisme de phase II consiste en la transformation des composés par des enzymes tels que les transférases de glucuronosyles (UGT) qui se retrouvent également majoritairement dans le foie. La clairance hépatique peut être déterminée en identifiant le temps de demi-vie du composé d'intérêt en l'incubant avec des microsomes hépatiques ou en utilisant des collections d'enzymes pures qui sont disponibles commercialement.

Dans le cas où les inhibiteurs afficheraient un mauvais profil d'absorption, réduire le nombre de cycles aromatiques retrouvés sur les inhibiteurs pourrait être bénéfique. De

récentes études suggèrent<sup>4,5,6</sup> que les composés possédant plus de trois cycles aromatiques<sup>3</sup> ou qui possèdent de nombreux carbones hybridés  $sp^2$  peuvent donner lieu à une faible absorption<sup>4</sup> dû à leur faible solubilité et leur lipophilicité élevée. De plus, ces composés ont plus tendance à inhiber les cytochromes P450<sup>3</sup>, les hERG<sup>3</sup> (canaux à potassium cardiaques) et à se lier avec l'albumine<sup>3</sup>. En effet, les médicaments actuellement utilisés possèdent en moyenne 1.6 cycle aromatique<sup>3</sup> tandis que le meilleur composé identifié lors de ce projet de recherche (composé **8**) en possède 6. Ainsi, il sera peut-être nécessaire de réduire le nombre de cycles aromatiques sur les inhibiteurs de la DHFR R67. De plus, en réduisant le nombre de cycles aromatiques, la solubilité des inhibiteurs pourrait augmenter ce qui permettrait d'estérifier les carboxylates et ainsi, permettre aux inhibiteurs de diffuser à travers la membrane plus facilement.

Ainsi, ce projet a permis d'identifier une gamme de molécules aux structures variées qui inhibent la DHFR R67 avec une affinité variant entre le millimolaire et les bas micromolaires. Plusieurs outils servant à l'identification et l'étude d'inhibiteurs de la DHFR R67 furent développés. Enfin, des pistes furent proposés afin de poursuivre l'amélioration de ces composés, en vue du développement d'inhibiteurs plus puissants qui pourront ultimement entrer dans le développement d'un antibiotique visant à briser la résistance au TMP conférée par la DHFR R67.

## Références du chapitre 6

1. Chopra, S.; Lynch, R.; Kim, S. H.; Jackson, M.; Howell, E. E., Effects of temperature and viscosity on R67 dihydrofolate reductase catalysis. *Biochemistry* **2006**, *45* (21), 6596-605.
2. Jackson, M.; Chopra, S.; Smiley, R. D.; Maynard, P. O.; Rosowsky, A.; London, R. E.; Levy, L.; Kalman, T. I.; Howell, E. E., Calorimetric studies of ligand binding in R67 dihydrofolate reductase. *Biochemistry* **2005**, *44* (37), 12420-33.
3. Park, H.; Bradrick, T. D.; Howell, E. E., A glutamine 67--> histidine mutation in homotetrameric R67 dihydrofolate reductase results in four mutations per single active site pore and causes substantial substrate and cofactor inhibition. *Protein Eng* **1997**, *10* (12), 1415-24.
4. Ritchie, T. J.; Macdonald, S. J., The impact of aromatic ring count on compound developability--are too many aromatic rings a liability in drug design? *Drug Discov Today* **2009**, *14* (21-22), 1011-20.
5. Lovering, F.; Bikker, J.; Humblet, C., Escape from flatland: increasing saturation as an approach to improving clinical success. *J Med Chem* **2009**, *52* (21), 6752-6.
6. Ritchie, T. J.; Macdonald, S. J.; Young, R. J.; Pickett, S. D., The impact of aromatic ring count on compound developability: further insights by examining carbo- and hetero-aromatic and -aliphatic ring types. *Drug Discov Today* **2011**, *16* (3-4), 164-71.

## **Annexe I - Matériel et Méthodes**

### **Chapitre 2**

#### **Section 2.2.1 - Solubilité des composés**

Les volumes correspondants de solvant organique ont été ajoutés à 50  $\mu$ M de DHF et NADPH en tampon phosphate de potassium 50mM à pH 7. Les cinétiques ont été démarrées par l'ajout de 200nM de DHFR R67. Les essais ont été préparés par une station de pipetage automatisé BioMek NX et les mesures d'absorbances ont été prises à 340nm par une lectrice à plaque Beckman DTX880. Les volumes des essais étaient de 200  $\mu$ L.

#### **Section 2.2.2 - Absorbance des composés**

Le composé **15** a été dissous dans le DMSO pur et les dilutions appropriées ont été faites en DMSO pur. Les dilutions du composé 15 ont été ajoutées à du tampon phosphate de potassium 50mM à pH 7 pour une concentration finale de 10% DMSO. Les mesures d'absorbance ont été prises à 340nm par une lectrice à plaque Beckman DTX880. Les volumes des essais étaient de 200  $\mu$ L et 100uL.

#### **Section 2.2.3 - pH du milieu réactionnel**

Les composés ont été dissous dans le DMSO pur et les dilutions appropriées ont été faites en DMSO pur. Les dilutions ont été ajoutées à du tampon phosphate de potassium 50mM à pH 7, contenant 50  $\mu$ M de DHF et de NADPH, pour une concentration finale de 10% DMSO. Les cinétiques ont été démarrées par l'ajout de 200nM de DHFR R67. Les essais ont été préparés par une station de pipetage automatisé BioMek NX et les mesures d'absorbances ont été prises à 340nm par une lectrice à plaque Beckman DTX880. Les volumes des essais étaient de 200  $\mu$ L ou de 100 uL dans le cas d'un composé qui avait une forte absorbance à 340nm. Le pH du milieu réactionnel a été mesuré avec le papier pH Hydrion provenant de FisherScientific.



### **Section 2.3 - Fiabilité de l'essai**

48 contrôles positifs et négatifs ont été préparés afin de déterminer le Zscore. 50  $\mu$ M de DHF et de NADPH ont été ajoutés à 50 mM de tampon phosphate de potassium à pH 7 contenant 10% de DMSO. Les réactions ont été démarrées par l'ajout de 200 nM de DHFR R67 (pour les contrôles positifs). Les essais ont été préparés par une station de pipetage automatisé BioMek NX et les mesures d'absorbances ont été prises à 340nm par une lectrice à plaque Beckman DTX880. Les volumes des essais étaient de 200  $\mu$ L.

## **Chapitre 3**

Voir la section matériel et méthode de l'article présenté à ce chapitre.

Précision à la suite des commentaires du jury de mémoire : La solubilité a été évaluée par inspection visuelle dans une éprouvette de verre.

## **Chapitre 4**

Tous les composés ont été synthétisés et fournis par le laboratoire du Dr. Jean-Jacques Vandenberghe (Université de Mons en Belgique). Les composés ont été dissous dans le DMSO pur et les dilutions appropriées ont été faites en DMSO pur. Les dilutions ont été ajoutées à du tampon phosphate de potassium 50mM à pH 7, contenant 50  $\mu$ M de DHF et de NADPH, pour une concentration finale de 10% DMSO. Les cinétiques ont été démarrées par l'ajout de 200nM de DHFR R67. Les essais ont été préparés par une station de pipetage automatisé BioMek NX et les mesures d'absorbance ont été prises à 340nm par une lectrice à plaque Beckman DTX880. Les volumes des essais étaient de 200  $\mu$ L ou de 100  $\mu$ L dans le cas d'un composé qui avait une forte absorbance à 340nm. Les régressions linéaires donnant la valeur de  $IC_{50}$  ont été faites avec le logiciel GraphPadPrism 5.00.

### **Section 4.5 - Arrimage moléculaire des composés 23 à 34**

Les molécules ont été dessinées avec ChemDraw Ultra 11.0. Une optimisation de la géométrie des structures a été faite avec HyperChem 8.0.3 avec l'algorithme Newton-Raphson selon les paramètres par défaut. L'arrimage moléculaire a été réalisé avec une

version modifiée du PDB 2RK1 de la DHFR R67 préparé par le Dr. Hervé Hogues (IRB-CNRC). Brièvement, l'orientation des Gln67 a été modifiée pour être cohérente avec la structure quaternaire de la DHFR R67 modélisée. Toutes les expériences d'arrimages moléculaires ont été réalisées avec le logiciel Moldock avec l'algorithme de recherche Moldock Optimizer et la fonction de score Moldock score [grid]. Les paramètres par défaut ont été utilisés excepté le nombre de poses créées qui a été modifié pour 5000 et le nombre de cycles qui a été modifié pour 20.

## Chapitre 5

### Section 5.1 - Criblage de la librairie virtuelle Maybridge

Toutes les expériences d'arrimages moléculaires ont été réalisées avec le logiciel Moldock avec l'algorithme de recherche est Moldock Optimizer et la fonction de score Moldock score [grid]. Les paramètres par défaut ont été utilisés excepté le nombre de poses créées qui a été modifié pour 2000 et le nombre de cycles qui a été modifié pour 5.

#### **Test *in vitro* de dix des composés Maybridge identifiés *in silico* et des composés du CNRC**

Les composés ont été dissous dans le DMSO pur et les dilutions appropriées ont été faites en DMSO pur. Les dilutions ont été ajoutées à du tampon phosphate de potassium 50mM à pH 7, contenant 50  $\mu$ M de DHF et de NADPH, pour une concentration finale de 10% DMSO. Les cinétiques ont été démarrées par l'ajout de 200nM de DHFR R67. Les essais ont été préparés par une station de pipetage automatisé BioMek NX et les mesures d'absorbances ont été prises à 340nm par une lectrice à plaque Beckman DTX880. Les volumes des essais étaient de 200  $\mu$ L ou de 100  $\mu$ L dans le cas d'un composé qui avait une forte absorbance à 340nm.